

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 2, 2005, 06:00:31 ; Search time 3385 Seconds  
(without alignments)  
9175.907 Million cell updates/sec

Title: US-10-720-018-1

Perfect score: 816  
Sequence: 1 atgacgactacattccac.....atggtgtcttactcttgg 816

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

```
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	276.6	33.9	899	9	CNS07770
2	129.2	15.8	969	6	CD457882 Fg04d_02e
3	118.4	14.5	682	4	BI750180
4	113.4	13.9	713	7	CF677471
5	113.4	13.9	836	7	CF701403
6	113.4	13.9	836	7	CF715948
7	112.8	13.8	777	7	CF817896
8	112.8	13.8	777	7	CF824542
9	111.2	13.6	738	7	CF812382
10	106.4	13.0	676	7	CF824543
11	102.8	12.6	718	1	AJ638396
12	100.8	12.4	615	7	CO136295
13	97.2	11.9	447	7	CO141255
14	93.2	11.4	511	1	AU249728
15	90.8	11.1	882	7	CN811863
16	82.8	10.1	758	7	CN812466
17	77.4	9.5	753	5	B0855471
18	77.2	9.5	896	6	CD456321
19	76.6	9.4	574	6	CD456321
20	76	9.3	833	7	CO012226
21	76	9.3	839	7	CO012227
22	76	9.3	922	7	CO009933
23	76	9.3	949	7	CO004456
24	76	9.3	963	7	CO004455

25	75.8	9.3	692	5	B092634
26	74.4	9.1	809	6	CB905184
27	74.4	9.1	809	7	CF876511
28	73.8	9.0	479	7	CF191935
29	73	8.9	816	2	BE642336
30	72.6	8.9	688	5	BP098973
31	72.6	8.9	735	7	CO140332
32	71.4	8.8	912	7	CO009934
33	70.8	8.7	690	7	CK569554
34	70.8	8.7	1048	9	CNS07622
35	70.2	8.6	712	6	CD458386
36	68	8.3	734	7	CF190316
37	68	8.3	734	7	CK447731
38	67.4	8.3	659	1	AJ638576
39	67.4	8.3	971	9	CNS07622
40	66.8	8.2	800	7	CO006699
41	65.4	8.0	786	5	B0855467
42	65.2	8.0	868	7	CO006700
43	64.6	7.9	700	7	CK567406
44	63.8	7.8	621	5	BU014570
45	63.6	7.8	703	5	B0861127

## ALIGNMENTS

RESULT 1	CNS07770	899 bp	DNA	1linear	GSS 08-JUL-2001
LOCUS	T7 end of clone B08A014G1 of library B08A from strain CBS 4732				
DEFINITION	of Pichia angusta, genomic survey sequence.				
ACCESSION	AL433186				
VERSION	AL433186.1				
KEYWORDS	GI:12216600				
SOURCE	GSS.				
ORGANISM	Pichia angusta				
REFERENCE	Bukaryovai; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Pichia.				
AUTHORS	1 (bases 1 to 899) Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durieux,P., Leplante,A., Lorente,B., Malpertuy,A., Neugebise,C., Ozier-Kalogiropoulos,O., Potier,S., Saurin,W., Tekai,F., Toffano-Nicolas,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.				
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies				
JOURNAL	FEBS Lett. 487 (1), 3-12 (2000)				
MEDLINE	20584711				
PUBMED	11152876				
REFERENCE	2 (bases 1 to 899) Blandin,G., Lorente,B., Malpertuy,A., Wincker,P., Artiguenave,F. and Dujon,B.				
AUTHORS	Genomic exploration of the hemiascomycetous yeasts: 13. Pichia angusta				
TITLE	FEBS Lett. 487 (1), 76-81 (2000)				
JOURNAL	20584723				
MEDLINE	11152888				
PUBMED	3 (bases 1 to 899)				
REFERENCE	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail : segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
AUTHORS	This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of				

FEATURES the other extremity of this insert.  
Location/Qualifiers

source

1..899

/organism="Pichia angusta"

/mol\_type="genomic DNA"

/strain="CBS 4732"

/db\_xref="taxon:4905"

/clone="BB0A014G11"

/clone\_1ib="BB0A"

/note="end : T7"

<18..>633

/note="similar to P50167 [ D-arabinitol 2-dehydrogenase

ABDH ] [ Pichia stipitis ]

1 putative frameshift(8)"

/evidence=not experimental

complement(<729..>812)

/note="similar to Saccharomyces cerevisiae ORF YKL126w [

YFK1 ; ser/chr-specific protein kinase ]"

/evidence=not experimental

complement(<732..>812)

/note="similar to Saccharomyces cerevisiae ORF YMR104c [

YPK2 ; ser/chr protein kinase ]"

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

/evidence=not experimental

RESULT 2

CD457882

LOCUS

DEFINITION

Fg04d\_02e10\_A

969 bp

mrna

linear

EST 14-JUN-2004

Accession

CD457882

VERSION

CD457882.1

KEYWORDS

EST

SOURCE

Gibberella zeae

ORGANISM

Gibberella zeae

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

Hypocryomycetidae; Hypocreales; Nectriaceae; Gibberella.

1 (bases 1 to 969)

De Moers, A., Hattori, J., Lacroix, C., Macott, M., Robert, L.S.,

Singh, J.A., Sprott, D. and Tinker, N.A.

Fusarium graminearum mycelium grown on wheat heads under high

humidity conditions

Unpublished (2003)

Contact: Ouellet, Therese

Eastern Cereal and Oilseed Research Centre

Agriculture and Agri-Food Canada

Neelby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA 0C6,

CANADA

Tel: (613) 759-1658

Fax: (613) 759-1701

Email: ouellet@agr.gc.ca.

Location/Qualifiers

1..969

/organism="Gibberella zeae"

/mol\_type="mrna"

/strain="DAOM 180378"

/db\_xref="taxon:5518"

/clone="Fg04d\_02e10"

/tissue\_type="Mycelium"

/dev\_stage="Asexual"

/lab\_host="E. coli"

/clone\_1ib="Fg04\_AAFc\_ECORC\_Fusarium\_graminearum\_mycelium\_

grown\_on\_wheat\_heads"

/note="Vector: Bluescript SK+, Site\_1: EcoRI; Site\_2:

XhoI; Fusarium grown on wheat (cv. Roblin) under high

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

humidity. CDNA made using Stratagene kit."

Query Match

Best Local Similarity

Matches

Conservative

3; Mismatches

206; Indels

8; Gaps

3;

Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;

15.8%; Score 129.2; DB 6; Length 969;

57.1%; Pred. No. 1.1e-25;



/note="Vector: pcwvSport6; Site\_1: NotI\_EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"

## ORIGIN

Query Match 13.9%; Score 113.4; DB 7; Length 713;  
Best Local Similarity 54.6%; Pred. No. 3.5e-21;  
Matches 278; Conservative 0; Mismatches 216; Indels 15; Gaps 2;

QY 308 TGGTTAACACAGCTGTTACTGTGAAACTTCCCATGTAAGATTACCCAGCAAGACG 367  
DB TCGTCACTGCTGCCGTATGTGCAAACTTGTGCTACAGATACCCCATGATTAAGA 590  
QY 368 CTGAGAGATGGTGAAGTTAACTTGTGCTTCTTGTATGTTCTCAAGCTTTGCTA 427  
DB TCAAGAGAGCTGTGGACATCAATATGGGTACTTGTGATGGCAGCTTGAAGCTGCA 530  
QY 428 AGCCATTGATCAAGAGGTATCAAGGCTTCTGTTGTTTGTATGTTGTTATGTCTG 487  
DB AGCTTATGCTGAA-----GGTGTTCATTACCTCGTCGATCTATAGAGG 482  
QY 488 GTGCAATTGTCAACCATCTCTCAAAACAGTTGTCTACAAAGATGTCCAGGCTGTGTA 547  
DB GTAGCATGTGTCAACGTTCTCTCAACCTCAAAACCTTTCAACCTTTTCAAGGCTGTGTC 422  
QY 548 TCCATTGTGCTAAGACTTGTGCTGTGTGATGGGCTTAAGTAAACATCAAGTTAATCTT 607  
DB GACACATGCTGATCTCTGCGCGTGTGATGGGCTCTCAAGGATCTCGTGTCAAGCTTC 362  
QY 421 GACACATGCTGATCTCTGCGCGTGTGATGGGCTCTCAAGGATCTCGTGTCAAGCTTC 362  
DB TAAACCCAGTTTACATCTACGCTCTTTGACCAAGATGTATCAATGTGAAGAAAT 667  
QY 361 TTAGTCCGGGTATACGCTTCAACCACTGATTAAGTCAATCTCGACGCCAACCCTTC 302  
DB TGTACAAAGATGATCTGTGTATCCCAACAAGAAATGTCCCAACCAAGAAATACA 727  
QY 668 TGTACAAAGATGATCTGTGTATCCCAACAAGAAATGTCCCAACCAAGAAATACA 727  
DB TCCGTGACAGTGTGTCAACCGTATCCCATGGGTGATGGCCGACCTTGTGATCTCA 242  
QY 728 TTGGTGTCTTTTGTACTTGTCTTGTGATGTGCTGCTTACATCACTATGTTGACGCT 787  
DB TGGTGTCTCTTCTTGTACTTGTCTTGTGATGTGCTGCTTACATCACTATGTTGACGCT 787  
QY 241 AGGGTCCGTCAATTACTTGTGCTTGA---CAGCTCAAGTACACCACTGATGCTGAGA 185  
DB TACTGTGTGATGATGTTTCACTTCTTGG 816  
QY 788 TACTGTGTGATGATGTTTCACTTCTTGG 816  
DB TCATGATTAACGCGGCTTACACTTGTCTG 156

## RESULT 5

CF701403/c 836 bp mRNA linear EST 16-AUG-2004  
LOCUS CCACCT16TF C.neoformans strain JEC21 Cryptococcus neoformans var.  
DEFINITION neoformans cDNA clone CCACCT16, mRNA sequence.  
ACCESSION CF701403  
VERSION CF701403.1 GI:41555562  
KEYWORDS EST.

ORGANISM Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)

REFERENCE Eukaryota; Fungi; Basidiomycota; Hymenomycetes;  
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;  
Filobasidiella.  
1 (bases 1 to 836)  
Loftus, B.

JOURNAL JEC21 cDNA library  
AUTHORS Unpublished (2003)  
TITLE End sequencing of clones from a full length enriched, normalized  
JEC21 cDNA library

COMMENT Other ESTs: CCACCT16TR  
TIGR Contact: Brendan Loftus

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208

Email: crypt@tigr.org  
Seq primer: TF.  
Location/Qualifiers  
1. 836

## FEATURES

source

/organism="Cryptococcus neoformans var. neoformans"  
/mol\_type="mRNA"  
/strain="JEC21"  
/db\_xref="taxon:40410"  
/clone="CCACCT16"  
/clone\_1ib="C.neoformans strain JEC21"  
/note="Vector: pcwvSport6; Site\_1: NotI\_EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"

## ORIGIN

Query Match 13.9%; Score 113.4; DB 7; Length 836;  
Best Local Similarity 54.6%; Pred. No. 3.7e-21;  
Matches 278; Conservative 0; Mismatches 216; Indels 15; Gaps 2;

QY 308 TGGTTAACACAGCTGTTACTGTGAAACTTCCCATGTAAGATTACCCAGCAAGACG 367  
DB TCGTCACTGCTGCCGTATGTGCAAACTTGTGCTACAGATACCCCATGATTAAGA 606  
QY 368 CTGAGAGATGGTGAAGTTAACTTGTGCTTCTTGTATGTTCTCAAGCTTTGCTA 427  
DB TCAAGAGAGCTGTGGACATCAATATGGGTACTTGTGATGGCAGCTTGAAGCTGCA 546  
QY 428 AGCCATTGATCAAGAGGTATCAAGGCTTCTGTTGTTTGTATGTTGTTATGTCTG 487  
DB AGCTTATGCTGAA-----GGTGTTCATTACCTCGTCGATCTATAGAGG 498  
QY 488 GTGCAATTGTCAACCATCTCTCAAAACAGTTGTCTACAAAGATGTCCAGGCTGTGTA 547  
DB GTAGCATGTGTCAACGTTCTCTCAACCTCAAAACCTTTCAACCTTTTCAAGGCTGTGTC 438  
QY 548 TCCATTGTGCTAAGACTTGTGCTGTGTGATGGGCTTAAGTAAACATCAAGTTAATCTT 607  
DB GACACATGCTGATCTCTGCGCGTGTGATGGGCTCTCAAGGATCTCGTGTCAAGCTTC 378  
QY 437 GACACATGCTGATCTCTGCGCGTGTGATGGGCTCTCAAGGATCTCGTGTCAAGCTTC 378  
DB TAAACCCAGTTTACATCTACGCTCTTTGACCAAGAAATGTATCAATGTGAAGAAAT 667  
QY 608 TAAACCCAGTTTACATCTACGCTCTTTGACCAAGAAATGTATCAATGTGAAGAAAT 667  
DB TTAGTCCGGGTATACGCTTCTCAACCACTGATTAAGTCAATCTCGACGCCAACCCTTC 318  
QY 377 TTAGTCCGGGTATACGCTTCTCAACCACTGATTAAGTCAATCTCGACGCCAACCCTTC 318  
DB TGTACAAAGATGATCTGTGTATCCCAACAAGAAATGTCCCAACCAAGAAATACA 727  
QY 668 TGTACAAAGATGATCTGTGTATCCCAACAAGAAATGTCCCAACCAAGAAATACA 727  
DB TCCGTGACAGTGTGTCAACCGTATCCCATGGGTGATGGCCGACCTTGTGATCTCA 258  
QY 728 TTGGTGTCTTTTGTACTTGTCTTGTGATGTGCTGCTTACATCACTATGTTGACGCT 787  
DB TGGTGTCTCTTCTTGTACTTGTCTTGTGATGTGCTGCTTACATCACTATGTTGACGCT 787  
QY 257 AGGGTCCGTCAATTACTTGTGCTTGTGACGCTCAAG---TACACCACTGATGCTGAGA 201  
DB TACTGTGTGATGATGTTTCACTTCTTGG 816  
QY 788 TACTGTGTGATGATGTTTCACTTCTTGG 816  
DB TCATGATTAACGCGGCTTACACTTGTCTG 172

## RESULT 6

CF715948/c 836 bp mRNA linear EST 16-AUG-2004  
LOCUS CCAB241TF C.neoformans strain JEC21 Cryptococcus neoformans var.  
DEFINITION neoformans cDNA clone CCAB241, mRNA sequence.  
ACCESSION CF715948  
VERSION CF715948.1 GI:41570107  
KEYWORDS EST.

ORGANISM Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)

REFERENCE Eukaryota; Fungi; Basidiomycota; Hymenomycetes;  
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;  
Filobasidiella.  
1 (bases 1 to 836)  
Loftus, B.

JOURNAL JEC21 cDNA library  
AUTHORS Unpublished (2003)  
TITLE End sequencing of clones from a full length enriched, normalized  
JEC21 cDNA library

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208

TITLE End sequencing of clones from a full length enriched, normalized JEC21 cDNA library  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: CCAB241TR  
Contact: Brendan Loftus  
TIGR  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208  
Email: crypt@tigr.org  
Seq primer: TF.

FEATURES  
source Location/Qualifiers  
1..836  
/organism="Cryptococcus neoformans var. neoformans"  
/mol\_type="mRNA"  
/strain="JEC21"  
/db\_xref="taxon:40410".  
/clone\_1lb="C.neoformans strain JEC21"  
/note="Vector: pCMVSPORT6; Site\_1: NotI EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"

# ORIGIN

Query Match 13.9%; Score 113.4; DB 7; Length 836;  
Best Local Similarity 54.6%; Pred. No. 3.7e-21;  
Matches 278; Conservative 0; Mismatches 216; Indels 15; Gaps 2;

308 TGGTAAACAGCTGGTACTGTGAAAGTCCCATGTGAGATTTCCCGCAAGAGAG 367  
670 TCTCTCTGCTGCGGATTTGCGAAACTTTCGCTACGAGTACCCCATGATGAGA 611  
368 CTGAGAGATGTGAGAGTTAATCTTGGTCTTGTATGTTCTCAAGCTTTGCTA 427  
610 TCAGAGAGCTGTGGACATCATATATGGTACTTGTATTCGCACTTGAGAGCTGCA 551  
428 AGCATTTGATCAAGAGAGTATCAAGGCTGCTTCTGTTTGTATGTTGTTATGTC 487  
550 AGCTTATGCTGAA-----GGTGTTCATTAACCTCGTGCATATAGAGCG 503  
488 GTGCCATTTGACATCTCCCAACCAAGTTCTCAACAATGCCAAGGCTGGTGA 547  
502 GTAGCATTTGCAAGTTCCTCAACCTCAACCCCTTCAACCTTCCAGGCTGCTG 443  
548 TCCATTTGGCTAAGCTTTGGCTTGTGAATGGCTAAGTCAATCAAGATTAATCTT 607  
442 GACACATGCTGATCCCTGCGCGTGAATGGCTCTCAAGGATATCCGTTCAAGCTC 383  
608 TAAACCAAGTTACATCTACGCTCTTGTGACCAAGATTTATCAATGTTACGAAAGAT 667  
382 TTAGTCCGGTTACGCTCTACCAACTGATTAAGTCAATTCGACGCAACCCGCTTC 323  
668 TGTAACAAGATGATCTCTGATCCCAACAAGAAATGTCGAAACCAAGATTAACA 727  
322 TCCGTACAGATGGCTCAACCGTATCCCATGSGTGAATGCGACCTTCTGATCTCA 263  
728 TTGGTCTGTTTGTACTTCTTCTGAATCTGCTCTTCAATCACTACTGTCAGAGCT 787  
262 AGAGTCCGCTCATTTACTTGTCTTGTGACAGCTCCAAG---TACACACATGCTGCTGAGA 206  
788 TACTGTTGATGTTGTTTCACTTCTTGG 816  
205 TCATGATGACGCGGTTACACTTGTCTTG 177

RESULT 7  
CF817896/c 777 bp mRNA linear EST 01-APR-2004  
LOCUS CF817896  
DEFINITION EST65278 Coccidioides posadasii saprobic phase cDNA library, greater than 4kb Coccidioides posadasii cDNA clone CIB147 5' end, mRNA sequence.

ACCESSION CF817896  
VERSION GI:45923774  
KEYWORDS EST.  
SOURCE Coccidioides posadasii  
ORGANISM Coccidioides posadasii  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides.

REFERENCE 1 (bases 1 to 777)  
Gardner M.J. and Cole G.T.  
Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags  
Unpublished (2003)  
Other ESTs: EST695277  
Contact: Gardner MJ  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301 838 3519  
Fax: 301 838 0208  
Email: gardner@tigr.org  
Seq primer: MJ3 Reverse.

FEATURES  
source Location/Qualifiers  
1..777  
/organism="Coccidioides posadasii"  
/mol\_type="mRNA"  
/strain="C735"  
/db\_xref="taxon:199306"  
/clone="CIB147"  
/dev\_stage="saprobic phase (mycelia)"  
/lab\_host="E. coli DH10B, TI phase resistant"  
/clone\_1lb="Coccidioides posadasii saprobic phase cDNA library, greater than 4kb"  
/note="Vector: pXpress 1; Site\_1: Not I; Site\_2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA > 4 kb"

# ORIGIN

Query Match 13.8%; Score 112.8; DB 7; Length 777;  
Best Local Similarity 54.1%; Pred. No. 5.3e-21;  
Matches 275; Conservative 0; Mismatches 227; Indels 6; Gaps 2;

304 CACTGGTTAACAAGCTGTTACTGTGAAAGTCCCATGTGAGATTTACCAGCAAG 363  
708 CATCTGTCATCTCGCGGTTCACTGAAACTTTGAGCGCGTCCATCCGATGAT 649  
364 AACGCTGAGAGATGTGAGGTTAATCTTGGCTTGTATGTTTCTCAAGCTTT 423  
648 CCATGCAAGAACTCGTCCGTCACAGTTGATGGTCAATCTTTGCACTGCTGT 589  
424 GCTAAGCATTAATCAAGAGATCAAGGCTGCTTCTGTTTGTATGTTTCTAG 483  
588 GCAAAACATCTAT--GGCTGCAAGTCCGTCGACGATTTGTTCAATGGAAGCATG 532  
484 TCTGTGCAATTTGCAACAGCTCTCAAAAACAGTTGTCTAACAATGTCGAAGCTGGT 543  
531 TCTGTGCAATTTGTAACGCTGCTCAAGCAAGCGCGTCAATGTCGAAGGCTGG 472  
544 GTTATCAATTTGCTAAGCTTTGCTGTAATGGGCTTAAGTACAACTCAAGATTAAT 603  
471 ATTAAGCACTTGGCTCTTCTTGGCAATGGAATGGGCAAGTGGGAATCAGAGTAAC 412  
604 TCTTTAACCAGTTTACATCTACGCTCTTGTGACCAAGATTTATCAATGTAACGA 663  
411 TGTATCAATGTCAGAGATCAATGTTGACTGCTGACCCGCAAGATTTGACGACAAACCT 352  
664 GAATTTACAAAGATGATCTCTGATCCCAACAAGATGTCGAACCAAGAGAA 723  
351 GATTTCAAGAGAGATGACCTCCCTCATTCACAAAGCAAAATGGGTACACTGAAGAC 292  
724 TACATTTGCTGTTTGTACTTCTTGTGATCTGCTTCAATCACTACTGCTGCC 783  
291 CTCATGAGCGCCAGTACGTTTATTTAGTGA---TGCAGCAAAATATGTTACTGGGCA 235  
784 AGCTTACTGTTGATGATGTTTCACTT 811

Db 234 GATTGAGGTTGATGCGGTACACTT 207

RESULT 8  
CF824542/c 787 bp mRNA linear EST 01-APR-2004  
LOCUS EST01924 Coccidioides posadasii saprobic phase cDNA library, 2 to  
DEFINITION 4 kb Coccidioides posadasii cDNA clone CIDAR90 3' end, mRNA  
sequence.

ACCESSION CF824542 GI:45930599  
VERSION CF824542.1  
KEYWORDS EST.  
SOURCE Coccidioides posadasii  
ORGANISM Coccidioides posadasii  
Bukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Onygenales; mitosporic Onygenales; Coccidioides.

REFERENCE 1 (bases 1 to 787)  
AUTHORS Gardner,M.J. and Cole,G.T.  
TITLE Analysis of gene expression in Coccidioides posadasii mycelia and  
spherules via expressed sequence tags  
JOURNAL Unpublished (2003)  
COMMENT Other\_ESTS: EST701925  
Contact: Gardner MJ  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301 838 3519  
Fax: 301 838 0208  
Email: gardner@tigr.org.

FEATURES  
source location/Qualifiers  
1..787

/organism="Coccidioides posadasii"  
/mol\_type="mRNA"  
/strain="C735"  
/db\_xref="taxon:199306"  
/clone="CIDAR90"  
/dev\_stage="saprobic phase (mycelia)"  
/lab\_host="E. coli DH10B, T1 phage resistant"  
/clone\_lib="Coccidioides posadasii saprobic phase cDNA  
library, 2 to 4 kb"  
/note="Vector: pEXpress 1; Site\_1: Not I; Site\_2: Eco RV;  
Coccidioides posadasii saprobic phase cDNA library, size  
fractionated cDNA 2 to 4 kb"

## ORIGIN

Query Match 13.8%; Score 112.8; DB 7; Length 787;  
Best Local Similarity 54.1%; Pred. No. 5.3e-21;  
Matches 275; Conservative 0; Mismatches 227; Indels 6; Gaps 2;

304 CACTTGTTTAAACAGCTGTTACTGTGAAACTTCCCATGTGAAGATTACCCAGCAAG 363  
|||  
706 CATCTGTCACATCTGCGGGTTCACTGAAACTTTGAGCGCGTGTCCACCGCATGAT 647  
364 AACCTGAGAGATGTTAAGTTAAGTTGCTTTGTGATGTTTCTCAAGCCTTT 423  
|||  
646 CCACTGACAAACTCTGCTCCGTCAAGTTGAGGTCATATCTTTTTCAGTTGCTGTT 587  
424 GCTAAGCCATTGATCAAGAAGTTAAGAGGTCCTGTTGTTTGTGATTTGATTTCTTANG 483  
|||  
586 GCGAAACATCTGAT--GGCTGCAAGTGCCTGGCAGCATTTGTTCATTGAAACATG 530  
484 TCTGTGCAATTTGTCACAGATCTTCAAAACCAAGTTTCTCAACATGTCGAAGCTGAT 543  
|||  
529 TCTGTGCAATTTGTCACAGATCTTCAAAACCAAGTTTCTCAACATGTCGAAGCTGAT 470  
544 GTTATTCATTTGCTTGAAGCTTTGCTTGAATGGGCTTAAGTACCAATCAAGATTAT 603  
|||  
469 ATTAAGGCACTTGCTGCTTCTTGGCAGTTGAATGGGCAAGTGGGAATCAGAATGAAC 410  
604 TCTTTAAACCCAGTGTACATCTAAGCTCTTTGACCAAGAATGTTATCAATGTATACGA 663  
409 TGTATCATGTCAGGATACATGTTGATGCTGCTGACCCGCAAGATCTTGTGACCAACCT 350

Qy 664 GAATTGTACAAACAGATGATCTGTGATCCACAAACAAAGATCTCGAACAAGGAA 723  
|||  
Db 349 GATCTCAAGAGAGATGAGACTCCCTCATTTCCAAAGGCAAAATGGGTACACCTGAAGAC 290  
Qy 724 TACATTTGCTGCTTTTGTACTGCTTTCTGATTTGCTGCTTCTTACACTACCTGTC 783  
Db 289 CTCATGGGCCCCAGTACACTTCTTATTTAGTGA---TGGAGCAAAATATGTTACTGGGCA 233  
Qy 784 AGCTTACTGTTGATGATGTTTCACTT 811  
Db 232 GATTGAGGTTTGAATGCGGTACACTT 205

RESULT 9  
CF812382/c 738 bp mRNA linear EST 01-APR-2004  
LOCUS EST689764 Coccidioides posadasii saprobic phase cDNA library,  
DEFINITION greater than 4kb Coccidioides posadasii cDNA clone CIBA394 3' end,  
mRNA sequence.

ACCESSION CF812382 GI:45918260  
VERSION CF812382  
KEYWORDS EST.  
SOURCE Coccidioides posadasii  
ORGANISM Coccidioides posadasii  
Bukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Onygenales; mitosporic Onygenales; Coccidioides.

REFERENCE 1 (bases 1 to 738)  
AUTHORS Gardner,M.J. and Cole,G.T.  
TITLE Analysis of gene expression in Coccidioides posadasii mycelia and  
spherules via expressed sequence tags  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST689765  
Contact: Gardner MJ  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301 838 3519  
Fax: 301 838 0208  
Email: gardner@tigr.org.

FEATURES  
source location/Qualifiers  
1..738

/organism="Coccidioides posadasii"  
/mol\_type="mRNA"  
/strain="C735"  
/db\_xref="taxon:199306"  
/clone="CIBA394"  
/dev\_stage="saprobic phase (mycelia)"  
/lab\_host="E. coli DH10B, T1 phage resistant"  
/clone\_lib="Coccidioides posadasii saprobic phase cDNA  
library, greater than 4kb"  
/note="Vector: pEXpress 1; Site\_1: Not I; Site\_2: Eco RV;  
Coccidioides posadasii saprobic phase cDNA library, size  
fractionated cDNA > 4 kb"

## ORIGIN

Query Match 13.6%; Score 111.2; DB 7; Length 738;  
Best Local Similarity 53.9%; Pred. No. 1.5e-20;  
Matches 274; Conservative 0; Mismatches 228; Indels 6; Gaps 2;

304 CACTTGTTTAAACAGCTGTTACTGTGAAACTTCCCATGTGAAGATTACCCAGCAAG 363  
|||  
612 CATCTGTCACATCTGCGGGTTCACTGAAACTTTGAGCGCGTGTCTTACCCCATGAT 553  
364 AACCTGAGAGATGTTAAGTTAAGTTGCTTTGTGATGTTTCTCAAGCCTTT 423  
|||  
552 CGCATGCAAAACTCTGCTCCGTCAAGTTATGAGGTCATATCTTTTTCAGTTGCTGTT 493  
424 GCTAAGCCATTGATCAAGAAGTATCAAGGATCTTCTGTTGTTTGAATTTGATTTANG 483  
|||  
492 GCGAAACATCTGAT--GGCTGCAAGTGCCTGGCAGCATTTGTTGATTTGAAGATG 436  
484 TCTGTGCAATTTGTCACAGATCTTCAAAACCAAGTTTCTCAACATGTCGAAGCTGAT 543  
435 TCTGTGCAATTTGTTAAGTGTCTCAGCCCAAGGCGCGTCAAAATGCTGCAAAAGTTGCG 376

QY 544 GTTATCCATTGGCTAAGACTTTGGCTTGGAATGGGCTTAAGTACAACATGAGTTAAT 603  
 DB 375 ATBAGGACCTTGCTCTCTCTCTGTCGATGGAATGGGCAATGTGGAAATCAAGTGAAC 316  
 QY 604 TCTTTTAAACCAAGTACATCTACGGCTCTTGAACCAAGATGTATTCATGCTTAACGAA 663  
 DB 315 TGTATCAGTCCAGGATACATGTTGACTGCCCTGACCCGCAAGATTCTTGACGACCAACCT 256  
 QY 664 GAATTGTACAACAGATGATCTCTGATATCCCAACAAAGATGTCGACCAACCAAGAA 723  
 DB 255 GATCTCAAGAGAGATGAGACCTCCATTCATCCCAAGGCAAAATGGGTACACTGAAGAC 196  
 QY 724 TACATGGGCTGTTTGTATGCTTCTTGATCTGCTGCTTCAATACACTACTGCTGCC 783  
 DB 195 CTCATGGGCGCCAGTGAAGTCTTATGAGAGA--TGCGACAAATATGTTACTTGAGGGA 139  
 QY 784 ACCTTACTGTTGATGCTGTTCACTT 811  
 DB 138 GATTGAGGTTGATGCGGCTACACTT 111

RESULT 10  
 CF824543 676 bp mRNA linear EST 01-APR-2004  
 LOCUS EST701925 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 Kb Coccidioides posadasii cDNA clone CIDAR90 5' end, mRNA sequence.

ACCESSION CF824543  
 VERSION CF824543.1 GI:45930600  
 KEYWORDS EST.  
 SOURCE Coccidioides posadasii  
 ORGANISM Coccidioides posadasii  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Oryzales; mitosporic Oryzales; Coccidioides.  
 1 (bases 1 to 676)  
 AUTHOR Gardner, M.J. and Cole, G.T.  
 TITLE Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags  
 JOURNAL Unpublished (2003)  
 COMMENT Other ESTs: EST701924  
 Contact: Gardner MJ  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301 838 3519  
 Fax: 301 838 0208  
 Email: gardner@ligr.org  
 Seq primer: M13 Reverse.

FEATURES  
 source Location/Qualifiers  
 1..676  
 /organism="Coccidioides posadasii"  
 /mol\_type="mRNA"  
 /strain="C735"  
 /db\_xref="taxon:199306"  
 /clone="CIDAR90"  
 /dev\_stage="saprobic phase (mycelia)"  
 /lab\_host="E. coli DH10B, T1 phage resistant"  
 /clone\_lib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"  
 /note="Vector: pEXpress 1; Site 1: Not I; Site 2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

ORIGIN  
 Query Match 13.0%; Score 106.4; DB 7; Length 676;  
 Best Local Similarity 58.6%; Pred. No. 3.6e-19;  
 Matches 204; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY 464 TGTGTTGATGTTGATGCTGAGTCCATTTGCAACGATCTCAAAACCAAGTTGCT 523  
 DB 121 TGTGTTGATGAGACATGCTGAGTCCATTTGTAACGAGCTCAGCCACCAAGCCCG 180  
 QY 524 ACAACATGTCGAAGGCTGTTATCATTTGGCTTAAGACTTTGGCTTGGAATGGGCTA 583

DB 181 ACAATCCCTCAAAAGCTGGGATTAAGGCACCTTGCTGCTCTTGAGATGATGGCAA 240  
 QY 584 AGTACAACATCAAGATTAATCTTAAACCAAGTTACATCTACGCTCTTGACCAAGA 643  
 DB 241 GTGGGGAATCAGAGTAACCTATACATGTCAGAGTACATGTTGACTGCTGACCCGGA 300  
 QY 644 ATGTTATCAATGTATAAGAAATTTGACAAAGATGATCTGTGATCCCAACAA 703  
 DB 301 AGATTCTTGACACAACCTGATCTCAAGAGAGGAGACCTCCATTCATTCACAAAGGA 360  
 QY 704 GATGTCGAACCAAGAAATATCATTTGCTGTTTGTACTGCTTTGTAATCTGCTG 763  
 DB 361 AAATGGTACACTTAAGACCTCATGGGCGCCAGTACGTTCTTATGACTGA--TGGA 417  
 QY 764 CTTCATACACTGCTGCGACCTTACTGTTGATGCTGTTCACTT 811  
 DB 418 GCAATATGTTACTGGGCGAGATTGAGGTTGATGGCGGTTACACTT 465

RESULT 11  
 AJ638396 718 bp mRNA linear EST 05-MAY-2004  
 LOCUS AJ638396 Mgc Mycosphaerella graminicola cDNA clone mgc0602f, mRNA sequence.

ACCESSION AJ638396  
 VERSION AJ638396.1 GI:47031453  
 KEYWORDS EST.  
 SOURCE Mycosphaerella graminicola  
 ORGANISM Mycosphaerella graminicola  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et Chaetothyriomycetes incertae sedis; Mycosphaerellaceae;  
 Mycosphaerella.  
 1 (bases 1 to 718)  
 AUTHOR Keon, J.P.R., Hargreaves, J.A., Antoniw, J.F. and Hammond-Kosack, K.  
 TITLE Analysis of expressed sequence tags from the wheat fungal leaf blotch pathogen, Mycosphaerella graminicola (anamorph Septoria tritici)  
 JOURNAL Fungal Genet. Biol. (2004) In press  
 COMMENT Contact: Keon J  
 Plant Pathogen Interactions Division,  
 Rothamsted Research,  
 Harpenden, Herts, UNITED KINGDOM  
 Tel: +44(0)1582 763133  
 Fax: +44(0)1582 760981  
 Email: john.keon@bbsrc.ac.uk  
 Insert Length: 800 Std Error: 100.00  
 Seq primer: M13 reverse.

FEATURES  
 source Location/Qualifiers  
 1..718  
 /organism="Mycosphaerella graminicola"  
 /mol\_type="mRNA"  
 /strain="Strit"  
 /db\_xref="taxon:54734"  
 /clone="mgc0602f"  
 /clone\_lib="Mgc"  
 /note="Vector: pSPORT1; library constructed from senescent wheat leaves 21-25 days after infection with Mycosphaerella graminicola exhibiting abundant hyphal growth and asexual sporulation"

ORIGIN  
 Query Match 12.6%; Score 102.8; DB 1; Length 718;  
 Best Local Similarity 53.0%; Pred. No. 4e-18;  
 Matches 268; Conservative 0; Mismatches 232; Indels 6; Gaps 2;

QY 304 CACTTGTTAACAACGCTGTTACTGTGAACCTTCCCATGTGAAGATTACCAAGCAAG 363  
 DB 49 CACTTGTTACTTCGCTGCTTCAACCGAAGACTTGACGCCATGACGATCCACAGCAG 108  
 QY 364 AACGCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 423  
 DB 109 CCGATCCAAACCTTTGGGAGTCAACGTCGACGACATATCATCTGCTTGACACCGTG 168



QY	424	CTAAGCATTGATCTAAGAAGATGTCAAAGGTCCTCGTTGTTGATATGGTCTAAG	483
Db	169	GCGAAGC---ATCTATGAGGCGCAAGGCTTCTGGACGCTGTGATGATGGTAGCAATG	225
QY	484	TCTGTGCATTGTCAAGATCCTCAAAACCAAGTTGTCACACATGTCACAAAGCTGAT	543
Db	226	TCTGGCGGCATGCTAATGTGCTCAGCCACAGGCTCCTTACACCGGGCCAAAGGTGCT	285
QY	544	GTTATCCATTGGCTAAGACTTTGCTGTGTGAATGGGCTAAGTACACATCAGAGTTAAT	603
Db	286	GTCGCGCATTCGACAGCTCGCTCCAGACGATGGGCTCAACGTGGCAATGTTGTCAAC	345
QY	604	TCTTTAAACCGAGTTACATCTACGGTCCTTTGACCAAGAAATGTTATCAATGTACGAA	663
Db	346	TGCATCTCTCTGGGCTACATGCTCACTGACTGACTACACAGAAAGATCCTTGACGACAAACCC	405
QY	664	GAATGTATCAACAGATGATCTCTGGATATCCACACAAAGAAGATGTCGACCAAGAA	723
Db	406	GAGCTCGCCAAAGTAGACAGGCTCATCCACAGGGCAAGTGGGTGGCCCAAGAGC	465
QY	724	TACATGTGCTGCTGTTTGTACTTGTCTGATCTGCTCTTACATACACTAAGTGTGCC	783
Db	466	CTGATGGGCGCGGTGACCTTCTCTCAGCGA---CGCGCGGAGTACGTGACCGGCGCC	522
QY	784	AGCTTACTGTTGATGATGATGTTTAC	809
Db	523	GATTCGCGTGTGACGGTGAGTACAC	548

FEATURES	source	Location/Qualifiers
LOCUS	COL136295	615 bp, mRNA, linear, EST 17-JUN-2004
DEFINITION	ES1830966 Aspergillus flavus Normalized cDNA Expression Library	
REFERENCE	Aspergillus flavus cDNA clone NAFBQ16 5' end similar to D-arabinitol 2-dehydrogenase (ribulose forming) (BC 1.1.1.250) (ABDH). [Yeast] Candida tropicalis, mRNA sequence.	
ACCESSION	CO136295	
VERSION	CO136295.1	GI:48885273
KEYWORDS	EST.	
SOURCE	Aspergillus flavus	
ORGANISM	Aspergillus flavus	
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.	
AUTHORS	Yu, J., Whitelaw, C.A., Nieman, W.C., Bhanagar, D. and Cleveland, T.E.	
TITLE	1 (bases 1 to 615)	
JOURNAL	Aspergillus flavus expressed sequence tags for identification of genes with putative roles in aflatoxin contamination of crops	
COMMENT	FEWS Microbiol. Lett. (2004) In press	
	Contact: Yu J	
	Food and Feed Safety Research Unit	
	USDA/ARS, Southern Regional Research Center	
	1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA	
	Tel.: 504 286 4405	
	Fax: 504 286 4419	
	Email: jlyu@rrc.ars.usda.gov	
	Contact Dr. Yu at USDA/ARS SRRC (jlyu@rrc.ars.usda.gov) for clone information	
	PCR Primers	
	FORWARD: M13F	
	BACKWARD: M13R	
	Seq primer: M13 Forward	
	POLYA=NO.	

```

FEATURES
  source
    location/Qualifiers
      1. .615
        /organism="Aspergillus flavus"
        /mol_type="mRNA"
        /strain="NRRL 3357"
        /db_xref="taxon:5059"
        /clone="NAFBO16"
        /sex="asexual mycelia"
        /cell_type="mycelia"
        /dev_stage="developmental stages from 18 to 96 hours"

```

/lab host: E. coli DH10B Thesistant cells"  
 /clone\_lib="Aspergillus flavus Normalized cDNA Expression  
 Library"  
 /note="Vector: pBluescript (SK+) (Stratagene), antibiotic  
 selection marker: Carbenicillin; Site 1: NotI, at the 5  
 prime end; Site 2: EcoRI, at the 3 prime end; This  
 normalized cDNA expression library was constructed using a  
 mixture of mycelial cells grown under eight different  
 medium conditions and harvested at 5 time points (18, 24,  
 48, 72, 96 hours). The poly-A sequence was trimmed off  
 before ligating to vector."

Query Match	12.4%	Pred. 100.8;	DB 7,	Length 615;
Best Local Similarity	51.2%	Pred. No. 1.4e-17,		
Matches 287; Conservative	0;	Mismatches 267;	Indels 6;	Gaps 22;

Oy	196	TTGAAAGAAAGTTCGAAAGATGGGTTCAATATCCCTGTGATATTTTCGATTCGATACCGTT	255
Db	38	TTGGAAACAAATGCCCCAAGCTCACCGCCCACTAACGCTATATGTTCCGACCCTTAACCTCGCT	97
Oy	256	CACAAGGTGTTGCTCAAGTTCCTAAGAGATTTTGGTAAAGTTGCCATTCGACTTGGTTAAC	315
Db	98	AACGATGCCCTCTCCGATATATATCTCCAAAGCAGGGCAAGAT---GACAAACTGGTCAACC	154
Oy	316	ACAGCTGTTACTGTGAAAACTTTCCCATGTGAAAGTTATCCAGCCAAACAGCTGAGAG	375
Db	155	TCCGCCGATTCACGGAAACCTTGATGCCATCTCTACCTCAACGACCGTCTGCMAAAG	214
Oy	376	ATGGTGAAGTTACTTGTTGGGTTCTTTGATGTTCTCAAGCCTTGTCTAAGCATTG	435
Db	215	CTTTGGGGCGTTAAAGTCGATGGAACAATCTTTTGGCACCGGATGTGCCAAGC---AC	271
Oy	436	ATCAAAGAGGTATCAAAGGTGCTTGTGTGTTTGAATGATGTTCTATGTCTGTGGTCC	495
Db	272	CTCATGAGGCGCAAGGTTCCGGGACGATTTGTCTATGATTTGTAGCATGTCTGTGGTATCC	331
Oy	496	GTCAACGATCTCTAAACCAACAGTTGTCTACAACAATGTCCAAAGCGTGGTGTATCCATTG	555
Db	332	GTCAACGTGCCGACGCCCAAGGCTCTTACCAACGCCGCGAAGGCGCGCTGTGCTCACTT	391
Oy	556	GCTAAGACTTTGGGCTTGATGATGGGCTTAAGTACAACAATCAGATTAATCTTTTAAACCA	615
Db	392	GCCGCGTCTTGGCGCTGGAATGAGCGCGGTCAACAATCCGGGTGAATCGATCAAGCCCT	451
Oy	616	GGTATACATCTACGGTCTTTGACCAAGATGTTATCATGAGTAAAGAAATTTGTACAAC	675
Db	452	GGATACATCTTACTGCCCCCTBACCCGCAAGATTTTGGATGAGAACCCGGAATTTGGGGAC	511
Oy	676	AGATGATCTCTGAGTATCCCAACAAGAAATGTCGAACCAAGAAATTAATTTGGTGTCT	735
Db	512	AAGTGAATCTGCTATATCCCAACCGGCAAGATGGATCTCCCGAGGACCTGATGGTCCC	571
Oy	736	GTTTTGTACTGCTTTCTGGA	755
Db	572	GTTAACCTTCCCTCATGGA	591

RESULT 13				
LOCUS	COI41255			
DEFINITION	COI41255	447 bp	mRNA	linear EST 17-JUN-2004
ACCESSION	EST835926	Aspergillus flavus	Normalized cDNA	expression library
VERSION	Aspergillus flavus	cDNA clone	NAFEA37.5	end, mRNA sequence.
KEYWORDS	COI41255.1	GI:48890246		
SOURCE	EST.			
ORGANISM	Aspergillus flavus			
REFERENCE	Aspergillus flavus			
AUTHORS	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
TITLE	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus			
	1 (bases 1 to 447)			
	Yu,J., Whitelaw,C.A., Nieman,W.C., Bhatnagar,D. and Cleveland,T.E			
	Aspergillus flavus expressed sequence tags for identification of			



JOURNAL  
COMMENT

genes with putative roles in aflatoxin contamination of crops  
 FEMS Microbiol. Lett. (2004) In press  
 Contact: Yu J  
 Food and Feed Safety Research Unit  
 USDA/ARS, Southern Regional Research Center  
 1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA  
 Tel: 504 286 4405  
 Fax: 504 286 4419  
 Email: jinyu@rrc.ars.usda.gov  
 Contact Dr. Yu at USDA/ARS SRRC (jinyu@rrc.ars.usda.gov) for clone  
 information  
 PCR Primers  
 FORWARD: M13F  
 BACKWARD: M13R  
 Seq primer: M13 Forward  
 POLYA=No.

FEATURES  
source

Location/Qualifiers  
 1..447  
 /organism="Aspergillus flavus"  
 /mol\_type="mRNA"  
 /strain="NRRL 3357"  
 /db\_xref="taxon:5059"  
 /clone="NAFEA37"  
 /sex="asexual mycelia"  
 /cell\_type="mycelia"  
 /dev\_stage="developmental stages from 18 to 96 hours"  
 /lab\_host="E. coli DH10B T1 resistant cells"  
 /clone\_lib="Aspergillus flavus Normalized cDNA Expression  
 library"  
 /note="Vector: pBluescript (SK+) (Stratagene), antibiotic  
 selection marker: Carbenicillin; Site 1: NotI, at the 5  
 prime end; Site 2: EcoRI, at the 3 prime end; This  
 normalized cDNA expression library was constructed using a  
 mixture of mycelial cells grown under eight different  
 medium conditions and harvested at 5 time points (18, 24,  
 48, 72, 96 hours). The poly-A sequence was trimmed off  
 before ligating to vector."

## ORIGIN

Query Match 11.9%; Score 97.2; DB 7; Length 447;  
 Best Local Similarity 58.7%; Pred. No. 1.4e-16;  
 Matches 168; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

464 TTGTTTGAATGTTATGTCGTGTCATTTGTCAGATCTCTCAAAACCAAGTGTCT 523  
 162 TTGTCATGATGTTAGCATGTCGTGTCATGTCACAGTCCGCGACGCCAGCTCTT 221  
 524 ACAACATGTCGAAGCTGCTGTCATTTGCTAAGACTTGGCTTGTAATGGGCTA 583  
 222 ACAAGCCCCGCAAGGCCGCTGTCGCACTTCCGCGTCCTTCCGCGCAATGGGCG 281  
 584 AGTACAACATCAGAGTTAATTTCTTAACCCAGGTTCATCTACGGTCTTTGACCAAG 643  
 282 GTACAGCATCCGGGGAAGTGCATCAGCCCTGATACATGCTTACTGCGCTGACCGCA 341  
 644 ATGTTATCATGTGTAACGAATTTGTACAAGATGATCTTGTGTAATCCCAACAA 703  
 342 AAGATTTTGAATGAAACCCGGAATTCGGGACAAAGTGAATCTGCTATCCCAACGGCA 401  
 704 GAATGCCGAACCAAGATATCATTTGCTGTTTGTACTGCT 749  
 402 AAGTGGTACTCTCCGAGACGCTGATGGTCCGTTACTTCTGCT 447

## RESULT 14

AU249728 511 bp mRNA linear EST 22-APR-2004  
 LOCUS AU249728 SL Lolium multiflorum cDNA clone SL002A09-5, mRNA  
 DEFINITION sequence.  
 ACCESSION AU249728  
 VERSION AU249728.1 GI:46506997  
 KEYWORDS EST.  
 SOURCE Lolium multiflorum (Italian ryegrass)

## ORGANISM

Lolium multiflorum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Pooideae; Poaeae; Lolium.  
 1 (bases 1 to 511)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Lolium multiflorum EST Project  
 Unpublished (2004)  
 Contact: Seishi Ikeda  
 Japan Grassland Farming Forage Seed Association (JFSA)  
 Forage Crop Research Institute (FCRI)  
 Higashiakada 388-5, Nishinauno, Tochigi 329-2742, Japan  
 Tel: 81-287-37-6755  
 Fax: 81-287-37-6757  
 Email: sikedas7@jfsaas.or.jp  
 Contact: Takashi Takamizo (takamizo@affrc.go.jp)  
 National Institute of Livestock and Grassland Science, Nishinasuno  
 Resistance gene analog.

FEATURES  
source

Location/Qualifiers  
 1..511  
 /organism="Lolium multiflorum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4521"  
 /clone="SL002A09-5"  
 /tissue\_type="Powdery mildew infected leaf"  
 /clone\_lib="SL"

## ORIGIN

Query Match 11.4%; Score 93.2; DB 1; Length 511;  
 Best Local Similarity 58.5%; Pred. No. 2.1e-15;  
 Matches 200; Conservative 0; Mismatches 138; Indels 4; Gaps 2;

469 TTGATGTTTCATGTCGTGTCATTTGTCAGATCTCTCAAAACCAAGTGTCTACAC 528  
 2 TTTATTTGAAGCATGTCAGAGACCGTCGTAATGTTCCAAACCAACCAACCATTAAC 61  
 529 ATGTCCAGAGCTGAGTTATTCATTTGCTAAGACTTGGCTTGTAATGGGCTAATAC 588  
 62 GCCCTAAGCTGCAATTTGGCACTGSCATGAGTTCAGATGGAGTGGGCTCATGCA 121  
 589 AACATCAGATTAATTTCTTAACCCAGGTTCATCTACAGTCTTTGACCAAGATGTT 648  
 122 GGAATTCAGATCAATTTGATTTGCTGATACATGCTACATGATGACCAAAAAAAT 181  
 649 A-TCAATGTAACGAAGATTTGTACAAGATGATCTTGTGTAATCCCAACAAAGAT 707  
 182 ACTCGAAGGAATCTGATTTAAGAGCAATGATCTCGTATTCACAAAGGCAAAAT 241  
 708 GTCCGAACCAAGGAATATATGTCGTGTTTGTACTGCTTTCGATCTGCTGCTC 767  
 242 GGTATATCCAGAGATTTGATGGAGCAATTAATCTTCTGCTCTGA--TGCCTCATC 298  
 768 ATACACTACTGTCGACGCTTACTGCTTGTGATGATGATTTTAC 809  
 299 ATACATGACGGGCGCTGATTAAGATGATGGGGCTACAC 340

## RESULT 15

CN811863/892 bp mRNA linear EST 01-JUN-2004  
 LOCUS CN811863 Fg09\_08p20\_A Fg09\_AARC ECORC Fusarium graminearum simple\_substrate  
 DEFINITION Gibberella zeae cDNA clone Fg09\_08p20, mRNA sequence.  
 ACCESSION CN811863  
 VERSION CN811863.1 GI:47835874  
 KEYWORDS EST.  
 SOURCE Gibberella zeae  
 ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
 1 (bases 1 to 892)  
 Watson, R.J., Heyes, R., Couroux, P., De Moors, A., Harris, L.J.,  
 Hattori, J., Lacroix, C., Masotti, M., Ouellet, T., Robert, L.S.,  
 Singh, J.A., Sprott, D. and Tinker, N.A.

TITLE A cDNA library prepared from *Fusarium graminearum* grown on a simple substrate  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Watson, Robert J.  
 Eastern Cereal and Oilseed Research Centre  
 Agriculture and Agri-Food Canada  
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
 CANADA  
 Tel: (613) 759-1655  
 Fax: (613) 759-1701  
 Email: watsonrj@agr.gc.ca

FEATURES  
 source  
 1. 892  
 Location/Qualifiers

/organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="DAOM 180378"  
 /db\_xref="taxon:5518"  
 /clone="Fg09\_08p20"  
 /issue\_type="Mycelium"  
 /dev\_stage="Asexual"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="Fg09\_AFC\_ECORC\_Fusarium\_graminearum\_simple\_substrac"  
 /note="Vector: pBluescript II+, Site\_1: EcoRI; Site\_2: XhoI; *Fusarium graminearum* grown on a simple substrate--minimal media supplemented with amino acids."

## ORIGIN

Query Match 11.1%; Score 90.8; DB 7; Length 892;  
 Best Local Similarity 54.5%; Pred. No. 1.2e-14;  
 Matches 204; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

```

OY 437 TCAAGAGAGGTATCAAGGGTCTTGTGTTTGTGTTTCTATGTCGAGCCATTG 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 388 TCATGAGAGGTCAAGGCCCGTAGTATGTGTATTTGAAGCATGTCTGTCATTG 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 497 TCAAGCATCTCAAAACCAAGTTGTCTACACATGTCGAAGCTGTGTATTCATTGG 556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 328 TCAATGTCCACAGCCAGCCCATATTAACGCGCAAAAGCAGCTGTCGACACTGG 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 557 CTAAGACTTGGCTTGTGAATGGGCTAATACATCATGAGTTAATCTTTAAACCCAG 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 268 CTGCTTCTTTCAGTGGAGTGGCTCAGCTGGAATCCGTGTCACTGCATCTCTCTG 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 617 GTTACATCTACGGTCCCTTGACCAAGATGTTATCAATGTTAACGAATTTGACAACA 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 208 GCTATATGTGTAAGTCTTTGACACAGAGATCTCGACGACAACTCTGATTTAGAGAGA 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 677 GATGATCTCTGTGATCCCAACAAGAAATGTCGAACCAAGAAATACATTTGTGCTG 736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 148 CCTGGACATCCCTCATTTCTCAGGGTGCATGGGACTGCTCAAGATTGATGGGACCCG 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 737 TTTTGTACTGTTTCTGTAATCTGTGCTTCACTACTCTGTGCAAGCTTACTGTGTTG 796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 88 TTAACCTTCTGCTATCAGA--TGGCTCACTTATATGACTGGGGCAAGTGTGAGTTG 32
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 797 ATGTGTGTTTCACT 810
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 31 ATGAGAGATACACT 18
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

Search completed: March 2, 2005, 16:54:05  
 Job time : 3394 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 2, 2005, 15:57:37 ; Search time 594 Seconds  
(without alignments)  
8140.026 Million cell updates/sec

Title: US-10-720-018-1

Perfect score: 816  
Sequence: 1 atgacgactacattccaac.....atgggtgttcaactctctg 816

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4394803 seqs, 296729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 4: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/prodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/prodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/prodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/prodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/prodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUBCOMB.seq:\*
- 22: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	816	100.0	816	18	US-10-720-018-1	Sequence 1, Appli
2	99.4	12.2	3203	18	US-10-653-047-41	Sequence 41, Appli
3	88.6	10.9	687	18	US-10-653-047-4932	Sequence 4932, Ap
4	81.8	10.0	637	17	US-10-369-493-36723	Sequence 36723, A
5	80.6	9.9	735	17	US-10-369-493-26991	Sequence 26991, A
6	74.6	9.1	852	10	US-09-734-237B-71	Sequence 71, Appli
7	74.2	9.1	1146	18	US-10-653-047-4825	Sequence 4825, Ap
8	72.4	8.9	775	18	US-10-425-115-172450	Sequence 172450,
9	67.6	7.8	783	17	US-10-369-493-25810	Sequence 25810, A
10	63.6	7.4	774	17	US-10-369-493-26610	Sequence 26610, A
11	60.6	7.4	636	18	US-10-653-047-3866	Sequence 3866, Ap

12	59.8	7.3	822	15	US-10-314-394-1	Sequence 1, Appli
13	56.4	6.9	825	15	US-10-314-394-3	Sequence 3, Appli
14	56.2	6.9	789	9	US-09-938-842A-948	Sequence 948, App
15	56.2	6.9	789	11	US-09-938-842A-948	Sequence 948, App
16	55.2	6.8	732	17	US-10-369-493-38959	Sequence 38959, A
17	55.2	6.8	732	17	US-10-369-493-38960	Sequence 38960, A
18	55.2	6.8	837	18	US-10-425-115-100705	Sequence 100705,
19	55	6.7	949	9	US-09-770-445-339	Sequence 339, App
20	54.4	6.7	783	9	US-09-938-842A-1025	Sequence 1025, Ap
21	54.4	6.7	783	11	US-09-938-842A-1025	Sequence 1025, Ap
22	53.8	6.6	768	17	US-10-369-493-26631	Sequence 26631, A
23	53.2	6.5	852	18	US-10-639-159-39	Sequence 39, Appli
24	53.2	6.5	852	18	US-10-639-159-53	Sequence 53, Appli
25	53.2	6.5	852	18	US-10-782-258-53	Sequence 53, Appli
26	53.2	6.5	855	18	US-10-639-159-45	Sequence 45, Appli
27	53.2	6.5	855	18	US-10-639-159-47	Sequence 47, Appli
28	53.2	6.5	855	18	US-10-782-258-39	Sequence 39, Appli
29	53.2	6.5	855	18	US-10-782-258-45	Sequence 45, Appli
30	53.2	6.5	855	18	US-10-782-258-47	Sequence 47, Appli
31	52.8	6.5	819	15	US-10-259-165-699	Sequence 699, App
32	52.4	6.4	656	9	US-09-770-149-497	Sequence 497, App
33	52.2	6.4	852	18	US-10-639-159-51	Sequence 51, Appli
34	52.2	6.4	852	18	US-10-782-258-51	Sequence 51, Appli
35	51.6	6.3	852	18	US-10-639-159-55	Sequence 55, Appli
36	51.6	6.3	852	18	US-10-639-159-57	Sequence 57, Appli
37	51.6	6.3	852	18	US-10-782-258-55	Sequence 55, Appli
38	51.6	6.3	852	18	US-10-782-258-57	Sequence 57, Appli
39	51.6	6.3	852	18	US-10-782-258-113	Sequence 113, App
40	51.6	6.3	852	18	US-10-782-258-119	Sequence 119, App
41	51.6	6.3	855	18	US-10-639-159-43	Sequence 43, Appli
42	51.6	6.3	855	18	US-10-782-258-43	Sequence 43, Appli
43	51.2	6.3	849	16	US-10-033-585-646	Sequence 646, Ap
44	51.2	6.3	711	15	US-10-259-165-770	Sequence 770, App
45	51.2	6.3	810	9	US-09-938-842A-1690	Sequence 1690, Ap

## ALIGNMENTS

RESULT 1  
US-10-720-018-1  
; Sequence 1, Application US/10720018  
; Publication No. US20040132074A1  
; GENERAL INFORMATION:  
; APPLICANT: Verho, Kltva  
; APPLICANT: Richard, Peter  
; APPLICANT: Penttila, Merja  
; TITLE OF INVENTION: New Enzyme for an in vivo and in vitro Utilisation of  
; TITLE OF INVENTION: Carbohydrates  
; FILE REFERENCE: 2530-120  
; CURRENT APPLICATION NUMBER: US/10/720, 018  
; CURRENT FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 816  
; TYPE: DNA  
; ORGANISM: Ambrosiozyma monospora  
US-10-720-018-1

Query Match	100.0%	Score 816;	DB 18;	Length 816;
Best Local Similarity	100.0%	Pred. No. 2.4e-230;		
Matches 816;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGACTGACTACATTCCATTGATTGATGGCCATTACATTGTCACAGTGCC	60	
DB	1	ATGACTGACTACATTCCATTGATTGATGGCCATTACATTGTCACAGTGCC	60	
QY	61	TGTGGTGGTTTACGCTGACCTTATCAAGGTTTGGCTTACGCTTGTGACATTGCT	120	
DB	61	TGTGGTGGTTTACGCTGACCTTATCAAGGTTTGGCTTACGCTTGTGACATTGCT	120	
QY	121	TTCCTGATATGACCAAAAAGACTGCTGCAAAACAAGCCGAATACCAAAATAGCT	180	

```

Db      121 TTGCTTGATATGACCCAAAGAAAAGACTGCTGCGCAACCAAGCGGAATATCCACAATATCGCT 180
Oy      181 ACTGAAGAATTGAAGTTGAAAAGAAGTTCCAAAAGATGGGTTTCATATGCTGTGATATTTCT 240
Db      181 ACTGAAGAATTGAAGTTGAAAAGAAGTTCCAAAAGATGGGTTTCATATGCTGTGATATTTCT 240
Oy      241 GATTCTGATACCGTTTCACAGAGGTTTGTCTCAAGTTGCTAAGATTTTGGTAAGTTGCCA 300
Db      301 TTGCACTTGTTAAACAACAGCTGGTTACTGTGAAAACCTTCCCATGTAAGATTACCCAGCC 360
Oy      361 TTGCACTTGTTAAACAACAGCTGGTTACTGTGAAAACCTTCCCATGTAAGATTACCCAGCC 360
Db      361 TTGCACTTGTTAAACAACAGCTGGTTACTGTGAAAACCTTCCCATGTAAGATTACCCAGCC 360
Oy      421 TTTGCTAAGCCATTGATCAAAAGAGGTATCAAAGGTCCTTGTGTTGTTTGAATGGTTCT 480
Db      421 TTTGCTAAGCCATTGATCAAAAGAGGTATCAAAGGTCCTTGTGTTGTTTGAATGGTTCT 480
Oy      481 ATGTCGTGGCCATTGTCTCAACGATCCTCAAAAACCAAGTTGTCTACAAACATGTCACAAAGGCT 540
Db      541 ATGTCGTGGCCATTGTCTCAACGATCCTCAAAAACCAAGTTGTCTACAAACATGTCACAAAGGCT 540
Oy      541 GGTGTTATTCATTTGGGCTPAAGACTTTGGCTGTGGAATGGGCTTAAGTACAACATCAGAGTT 600
Db      601 GGTGTTATTCATTTGGGCTPAAGACTTTGGCTGTGGAATGGGCTTAAGTACAACATCAGAGTT 600
Oy      601 AATTCCTTAAACCCAGGTTTACATCTACGGTCTTTGACCAAGAATGTTATCAATGGTAAAC 660
Db      601 AATTCCTTAAACCCAGGTTTACATCTACGGTCTTTGACCAAGAATGTTATCAATGGTAAAC 660
Oy      661 GAAGAAATGTATACAACAGATGATCTCTGGTATCCCAACAAGAAATGTCGGAACCAAG 720
Db      661 GAAGAAATGTATACAACAGATGATCTCTGGTATCCCAACAAGAAATGTCGGAACCAAG 720
Oy      721 GAATACATTTGCTGCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 780
Db      721 GAATACATTTGCTGCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 780
Oy      781 GCCAGCTTACTGTTGATGATGATGTTCACTTCTTGG 816
Db      781 GCCAGCTTACTGTTGATGATGATGTTCACTTCTTGG 816

RESULT 2
US-10-653-047-41
; Sequence 41, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groch Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 3203
; TYPE: DNA

```

```

; ORGANISM: Fusarium venenatum
;
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: (1) ..(3203)
;
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-41

```

Query Match	12.2%	Score 99.4;	DB 18;	Length 3203;
Best Local Similarity	52.7%	Pred. No. 4.3e-18;		
Matches 265; Conservative	0;	Mismatches 231;	Indels 7;	Gaps 2

Qy	308	TGGTTAACACAGCTGGTGTACTGTTGAAAACCTTCCCATGTGAAGATTACCAAGCAAGAAC	367
Db	2614	TAGTCACGTCCGGCTGCTTCAACGAGAAATTTGAGGCAATCAACTATCCATCATGATCGTA	2677
Qy	368	CTGAAAGAGTGTGAAGGTAACTTGTGGGTTCTTTGTATGTTTCTCAAGCTTTGCTA	427
Db	2674	TGCCCAAGTTGTGGGGTGTATATGTGACGGTACTTACCTCTTTGCAATTGCAATGGCCA	2733
Qy	428	AGCATTGATCTAAGAAGTATCAAGGGTGTCTGTGTTTGAATGGTTCTATGTCG	487
Db	2734	A--ACATCTATGGAAGCCAGGTACCTGTGGATGATGTGGTATATGGAAGCATGTCTG	2790
Qy	488	GTGCCATTGTCAACGATCCTCAAAAACAAGTTGTCTACAAATGTCCAAGGCTGGTGTTA	547
Db	2791	GTGCTATTGTCAATGTTTCCAAAGCCACAGCAACCATATAAAGCGGCCAAGGCACTGTTC	2855
Qy	548	TCCATTGGCTAAGACTTTGGCTTGTGATGGGCTTAAGTACAACTCAAGTTAAATCTT	607
Db	2851	GTCACTGGCTGTCTCCCTCCGAGTAGAGTGGGCTCAAGCTGGAATCCGAGTCAACTGT	2911
Qy	608	TAAACCAAGTTAATCTAATCTAAGGTCCTTTGACCAAGATGTTATCAATGTTAAGAAAT	667
Db	2911	TCTCTCTGGGTACATGTTGACTGTCTTAAAGCAGAAATTTGAAAGATTAACCAAGATA	2970
Qy	668	TGTACAAACAGATGATCTCTGATATCCCAACAAGAAATGTCGGAACCAAGAAATACA	727
Db	2971	TCGAGAGGACCTGAGACATCTTATATCTTCCAGGACGATAGGACTACTCAAGATCTGA	3033
Qy	728	TTGGTGCTGTTTTGTACTTCTTTTGAAATGTCTGCTTCAATPACATGACTGGTGCAGCT	787
Db	3031	TGGGTCTCTGT-----GACCTTTTGTGTGAGATGGGCTTCTTACATGACTGGGCGAGATC	3088
Qy	788	TACTGTTGAATGATGTTTACT	810
Db	3087	TTAGAGTAGATGAGAGATATACT	3109
RESULT 3			
US-10-653-047-4932			
; Sequence 4932, Application US/10653047			
; Publication No. US20040229367A1			
; GENERAL INFORMATION:			
; APPLICANT: Randy M. Berka			
; APPLICANT: Michael W. Rey			
; APPLICANT: Jeffrey R. Shuster			
; APPLICANT: Sakari Kauppinen			
; APPLICANT: Ib Groth Clausen			
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene			
; FILE REFERENCE: 5849.200-US			
; CURRENT APPLICATION NUMBER: US/10/653,047			
; PRIOR FILING DATE: 2003-08-29			
; PRIOR APPLICATION NUMBER: US/09/533,559			
; PRIOR FILING DATE: 2000-03-22			
; PRIOR APPLICATION NUMBER: 09/273,623			
; NUMBER OF SEQ ID NOS: 7860			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 4932			
; LENGTH: 687			
; TYPE: DNA			

ORGANISM: Aspergillus oryzae  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)...(687)  
 OTHER INFORMATION: n = A,T,C or G  
 US-10-653-047-4932

Query Match 10.9%; Score 88.6; DB 18; Length 687;  
 Best Local Similarity 49.8%; Pred. No. 2.8e-15;  
 Matches 279; Conservative 0; Mismatches 275; Indels 6; Gaps 2;

196 TTGAAGAAGTTCAGAAAGATGGGTCATATGCTGTGATATTTCTGATTTGATACCGTT 255  
 DB 114 TTGAACAATATGCCAAGTCACCGGCCACTAGCGATGTTTCCAGCCTAATCCGTC 173  
 QY 256 CACAAAGTGTGCTCAAGTGTCAAGATTTGGTAAGTGGCATTTGCACTTTGTTAAC 315  
 DB 174 AACGATGCCCTCCGATATTAATCTCAAGCAGCGCAAGAT---CGACAACCTGTACCC 230  
 QY 316 ACAGCTGTACTGTGTAAGAACTCCATGTGAAGATTACCCAGCAAGACGCTGAAG 375  
 DB 231 TCCGCGGATTCACGGAACCTTCATGCCATCTCCATCCCTCAGACCGCTGCAAAAG 290  
 QY 376 ATGTAAGTAACTTGTGGGTCCTTTGTATGTTTCAAGCCTTTGTAAGCATTTG 435  
 DB 291 CTTTGGGGGCTTAATATCATATGAAATACCTTTTCCAGCGGTGTCGCAAGC---AC 347  
 QY 436 ATCAAGAAAGTATCAAGGGTCTTGTGTTTGTATGTTTCAAGCTTTGTAAGCAT 495  
 DB 348 CTCATGAGACGCAAGGTCGCGGACGATTTGTATGTTTGTATGTTTGTATGTTTATC 407  
 QY 496 GTCAAGATCTTCAAAACCAAGTGTCTATCAATATCCAAAGCTGTGTATTCATTG 555  
 DB 408 GACAAAGTCCCGAGCCGACGAGCTCTTCAAGCGCCGCAAGGCTGTCTCAACTT 467  
 QY 556 GCTAAGACTTTGGCTTGTGATGGCTAAGTCAACATCGATGATTAATCTTTAAACCA 615  
 DB 468 GCCGCTACTTCCGCGGCAATGGCGGTCACGACATCCGGGTGAGCTGATGAGCCCT 527  
 QY 616 GGTACATCTACGCTCTTTGACCAAGATGTTATCAATGTAAGCAAGATTTGACAC 675  
 DB 528 CGATACATGCTTACCTGCGCTGACCCGCAAGATTTTGGATGAGAACCCCGATGGCGGAC 587  
 QY 676 AGATGATCTCTGTATTCACAAACAAGATGTCGACCAAGCAAGATATCATTTGCT 735  
 DB 588 AAGTGATCTCGCTCATCCCGCAAGAGTGTGATCTCCGAGGACCTGATGGGTCC 647  
 QY 736 GTTTGACTGCTTTCTGA 755  
 DB 648 GTTACCTTTCTGCTCAGNGA 667

RESULT 4  
 US-10-369-493-36723  
 Sequence 36723; Application US/10369493  
 Publication No. US2003023675A1

GENERAL INFORMATION:  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Hinkle, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Chen, Xianfeng  
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 FILE REFERENCE: 38-10(52052)B  
 CURRENT APPLICATION NUMBER: US/10/369,493  
 CURRENT FILING DATE: 2003-02-28  
 PRIOR APPLICATION NUMBER: US 60/360,039  
 PRIOR FILING DATE: 2002-02-21  
 NUMBER OF SEQ ID NOS: 47374  
 SEQ ID NO 36723  
 LENGTH: 637  
 TYPE: DNA

ORGANISM: Aspergillus nidulans  
 US-10-369-493-36723

Query Match 10.0%; Score 81.8; DB 17; Length 637;  
 Best Local Similarity 55.0%; Pred. No. 2.7e-13;  
 Matches 183; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

QY 304 CACTTGTTAACAAGCTGTGTTACTGTGAAACTTCCCATGTGAAGATTACCAAGCAAG 363  
 DB 70 CATCTGTACGTCCTGCGGAAATTTACGAAACTTGAAGCATACGATACCATACGAC 129  
 QY 364 AACGTGAGAGATGTGAAGTTAACTTTGGGTTCTTTGTATGTTTCAAGCTTT 423  
 DB 130 CGATGCAAGACTCTGCGGGGGTGAACGTGAACGCAATATCTGTTTCTACGGGCGTG 189  
 QY 424 GCTAAGCATTTGATCAAGAGATCAAGGGTCTTCTGTTGTTTGAATGGTTTATG 483  
 DB 190 GCAAGCACTTAT---GGAGCGCAAGGCGCCGGAGATATGTATGTATGATGATG 246  
 QY 484 TCTGTCATTTGTCAAGATCTCAAAACCAAGTTGTCTAACAATGTCAAGGCTGT 543  
 DB 247 TCTGTCATTTGTCAAGATCTCAAAACCAAGGCTCTCAACAGCGCCGCAAGGCGCC 306  
 QY 544 GTTATCATTTGGCTAAGACTTTGGCTTGTGATGGGCTTAAGTACAAATCAAGATTAT 603  
 DB 307 GTTCCGACCTTGTCTTCTGCTTCTTCCGCGCAATGGGAGAGATGTATCAAGGTTAAC 366  
 QY 604 TCTTTAAACCAAGTTACTCAAGGCTCTTTG 636  
 DB 367 TGCATTAAGCCGTGATACATGCTGACTCTG 399

RESULT 5  
 US-10-369-493-26991  
 Sequence 26991; Application US/10369493  
 Publication No. US2003023675A1

GENERAL INFORMATION:  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Hinkle, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Chen, Xianfeng  
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 FILE REFERENCE: 38-10(52052)B  
 CURRENT APPLICATION NUMBER: US/10/369,493  
 CURRENT FILING DATE: 2003-02-28  
 PRIOR APPLICATION NUMBER: US 60/360,039  
 PRIOR FILING DATE: 2002-02-21  
 NUMBER OF SEQ ID NOS: 47374  
 SEQ ID NO 26991  
 LENGTH: 735  
 TYPE: DNA  
 ORGANISM: Neurospora crassa  
 US-10-369-493-26991

Query Match 9.9%; Score 80.6; DB 17; Length 735;  
 Best Local Similarity 61.8%; Pred. No. 6.7e-13;  
 Matches 128; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 436 ATCAAGAAAGTATCAAGGTCCTTGTGTTTGTATGTTGTTATGTCGTCGCTT 495  
 DB 364 ATGATGCGCGTGGCAACCCGCTCATTAATCTTGTGCTCGATGTCGCAACATC 423  
 QY 496 GTCAAGATCTCAAAACCAAGTTGTCAACAATGTCGAAGCTGTGTTATTCATTG 555  
 DB 424 GTCAATACCTCTCAGAGAGATGTCGTTCAACAGCCTCTTAAGCTGTGTTATGACGTC 483  
 QY 556 GCTAAGACTTTGGCTTGTGATGGGCTAAGTACATCATGAGTTATTTTAAACCA 615  
 DB 484 GGCAGATCTTGTGCTGCGAGTGGGCAAGTACATCCTGTTAACTGCAATTTCCCC 543  
 QY 616 GGTACATCTAAGGTCCTTTGACCAAG 642







```

? SEQ ID NO 3866
? LENGTH: 636
? TYPE: DNA
? ORGANISM: Aspergillus niger
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)...(636)
? OTHER INFORMATION: n = A,T,C or G
US-10-653-047-3866

```

Query Match	7.4%	Score 60.6;	DB 18;	Length 636;
Best Local Similarity	51.3%	Pred. No. 5e-07;		
Matches 194; Conservative	0;	Mismatches 175;	Indels 9;	Gaps 2;

Qy	211	AAATGGGATCAATATCCGTGATAATTTCTGATTCGATACCGTTCAAGAGTGTTC	270
Db	200	AAATGACCTGCCATTAAGCTGATGTCCGAATCAAGAGATATTGAGAGGTTATCCAG	259
Qy	271	CAAGTTGCTAAGAGATTTTGGTAAGTTGCCAATGCACTTGTTAAACAAGCTGTTACTG	330
Db	260	CAGATTGCTTCAAGACTTTGGCAAGCTGTATATTATCGTGGTAAC--TCGGGGGTCACT	316
Qy	331	GAAAACTTCCCATGTATAAATTACCAGCCAAAGACGTGAAAGATGTGGAGTTAAC	380
Db	317	TCCAAATATGCGGCAGACAAAGACTTACACACGAAACATGCGCTGACATCAAGAAAGTCAAT	376
Qy	391	TTGTTGGGATTCCTTGATGATTTTCTCAAGCTTGTGTAAGCAATTGATCAAGAAAGATATC	450
Db	377	CTGATATGGCGCATTTTATATACCGCGCAAGGGGCTTCTCGGATCTTCAAGCAACAAGG---	432
Qy	451	AAGGTCCTTCCTGTTGTTTGAATGGATCTAATGTCGTGTGCATGTGTCAAAGATCTCAA	510
Db	433	--AATATGGAATGTATCTTCAACAGCTCTGTCACTGACAAATTTGGTGAATGTGCCCTCAG	480
Qy	511	AACCAAGTTGTCTACAACATGTCCAAAGCTGTGTATATCAATTTGGCTAAAGACTTTGGCT	570
Db	491	AAACAAGCTGCATAACAATGCTCTCAAGGTGGGGTCGTGCAAAATGGCAAAATGTCTCT	550
Qy	571	TGTGATATGGGCTAAGTAC	588
Db	551	GTGTGAATGGGTGATNTTC	568

```

RESULT 12
US-10-314-394-1
; Sequence 1, Application US/10314394
; Publication No. US20030143700A1
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Hiroaki
; APPLICANT: Ueda, Momoko
; APPLICANT: Pan, Ritsuzui
; APPLICANT: Hamatani, Takeshi
; TITLE OF INVENTION: METHODS FOR PRODUCING OPTICALLY ACTIVE ALCOHOLS
; FILE REFERENCE: SH2-006
; CURRENT APPLICATION NUMBER: US/10/314,394
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/385,434
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: JP 2002-152955
; PRIOR FILING DATE: 2002-05-27
; PRIOR APPLICATION NUMBER: JP 2001-375041
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Datura stramonium
; US-10-314-394-1

```

Query Match	7.3%	Score	59.8	DB	15	Length	822
Best Local Similarity	52.5%	Pred.No.	1e-06				
Matches	158	Conservative	0	Mismatches	137	Indels	6
						Gaps	1

OY	513	TCGCTCAACAGTCCCAAGCGTGGTATCATCTTTGGCTTAAGCTTTGGCTGTGAATG	578
OY	507	TCCTTACTAGCTTCCAAAGGTGCAATTAATCAATGAACAAAGGCTTTGGCTGTGATG	566
OY	579	GGCTAAGTACAAATCATGAGTTAATCTTTAAACCCAGGTACATCTTACGGTCTTGAAC	638
Db	567	GGCTTAAGACACATCTCGGTCAATTCAGTTGCTCCGGAATCATTTTAACCCACATGTT	626
OY	639	CAAGAATGTTATC-----AATGTACGAAGATTTGACACAGATGATCTCTGAT	692
Db	627	TGAACTGCAATTAAGAAAAATCTCATCAAAAAGAAATAGACAATTTTATTGCA	686
OY	693	CCACACAACAAAGATGTCCGAACCAAAGAAATACATTGTGTCTTTTGTACTGCTTTC	752
Db	687	GACTCTTAATGGGCGGGCCGGAAAGCCCCAAGAAATTTCTGCATTAATAGCTTTTCTTTC	746
OY	753	TGAATCGTGTCTTCATACACTACTAGTGTGCCAGCTTACTGTGTGATGTGTGGTTCACTTC	812
Db	747	CTTCCCTGCTTCATATATATTACGGGCGCAGATCATATGGGCTGACGGTGGATTCACAGC	806
OY	813	T 813	
Db	807	T 807	

```

RESULT 13
US-10-314-394-3
Sequence 3, Application US/10314394
Publication NO. US20030143700A1
GENERAL INFORMATION:
APPLICANT: Yamamoto, Hiroaki
APPLICANT: Ueda, Momoko
APPLICANT: Pan, Ritsuzui
APPLICANT: Hamatani, Takeshi
TITLE OR INVENTION: METHODS FOR PRODUCING OPTICALLY ACTIVE ALCOHOLS
FILE REFERENCE: SH2-006
CURRENT APPLICATION NUMBER: US/10/314,394
CURRENT FILING DATE: 2002-12-06
PRIOR APPLICATION NUMBER: US 60/385,434
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: JP 2002-152955
PRIOR FILING DATE: 2002-05-27
PRIOR APPLICATION NUMBER: JP 2001-375041
PRIOR FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 825
TYPES: DNA
ORGANISM: Hyoscyamus niger
US-10-314-394-3

```

Query Match	6.9%	Score 56.4	DB 15	Length 825
Best Local Similarity	52.4%	Pred. No. 1e-05		
Matches 151, Conservative	0	Mismatches 131	Indels 6	Gaps 1

Qy	532	TCCAAGCGTGGTATATCCATTTGGCTTAAGCTTGGCTGTGAATGGGCTAATGTCAC	591
Db	523	TCCAAGCTGCATATATCAATTAACGAAGACTTGGCATGTGATGGGCCAAGACAC	582
Qy	592	ATCAGATTAAATTCCTTTAAACCAGGTACATCTACGGTCTTGTACCAAGAAATGTTATC	651
Db	583	ATTGGGTCAATTCAGTTGCTCCAGAGTCAATTTTAAACCCACATCTGAACCTGCAATT	642
Qy	652	-----AATGTACAGAGAAATTGTACACAGATGATCTCTGTATCCCAACAAAGA	705
Db	643	AGAAAAATCTCATCTAAAAAGAGAAATAGACAATTTAATTCAGACATCCCAATGGGC	702
Qy	706	ATGTCCGATACCAAGGATATCACTGGTGTCTTTTGTACATTGCTTTCGAATCGGTGCT	765
Db	703	CGGCGTGAAGACCCCAATGAAGTGTCTCACTAAATAGCTTTCTTCTCTCCCTGCTGCT	762

QY 766 TCATACACTACTGCTGCAGCTTACTGTTGATGATGTTTCACTTCT 813  
DB 763 TCTTATATTAATCTGGCCAAATTTATATGCGCTGATGATGATTCACAGCT 810

## RESULT 14

US-09-938-842A-948  
; Sequence 948, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: S001300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 948  
; LENGTH: 789  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-948

Query Match 6.9%; Score 56.2; DB 9; Length 789;  
Best Local Similarity 55.2%; Pred. No. 1.1e-05;  
Matches 153; Conservative 0; Mismatches 118; Indels 6; Gaps 2;

QY 533 CCAAGGCTGCTGTTATCCATTTGGCTAAGACTTGGCTTGTGATGGGCTAATACACA 592  
DB 485 CAAAGAGACTTTGATTCAGCTAGCTTAATAATTTGGCATGTGATGGCAAAAGCGCA 544  
QY 593 TCAGGTTAATCTTTAAACCGATTAATCAATCAAGCTTCTTGAACCAAGATTTATCA 652  
DB 545 TAAGAGCCAAAGCTGTTGGCTTAATGATCAATCAATCTCTGTCTCAATCTTATCTTG 604  
QY 653 ATGCTAACGAATTTGATCAACAGATGATCTCTGTGATCCCAACAAAGATGTCG 712  
DB 605 AGG---ACGTCACTTCAAGAAAGCAATGTTGATGAGACTCCACTTGTGCTGTGAG 661  
QY 713 AACCAAGAAATCAATGCTGCTTTTGTGATCTGTTCTGAATGCTGCTTCAATACA 772  
DB 662 AGCCAAATGAG---TTGATCACTAGTGGCTTCTTGTCTTCACTGACGCTTCTTATA 718  
QY 773 CTACTGTCACGCTTACTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 809  
DB 719 TTACTGTCAGACTAATTTGTGTTGATGATGATGATGATGATGATGATGATGATGAT 755

## RESULT 15

US-09-938-842A-948  
; Sequence 948, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: S001300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 948  
; LENGTH: 789  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-948

Query Match 6.9%; Score 56.2; DB 11; Length 789;  
Best Local Similarity 55.2%; Pred. No. 1.1e-05;  
Matches 153; Conservative 0; Mismatches 118; Indels 6; Gaps 2;

QY 533 CCAAGGCTGCTGTTATCCATTTGGCTAAGACTTGGCTTGTGATGGGCTAATACACA 592  
DB 485 CAAAGAGACTTTGATTCAGCTAGCTTAATAATTTGGCATGTGATGGCAAAAGCGCA 544  
QY 593 TCAGGTTAATCTTTAAACCGATTAATCAATCAAGCTTCTTGAACCAAGATTTATCA 652  
DB 545 TAAGAGCCAAAGCTGTTGGCTTAATGATCAATCAATCTCTGTCTCAATCTTATCTTG 604  
QY 653 ATGCTAACGAATTTGATCAACAGATGATCTCTGTGATCCCAACAAAGATGTCG 712  
DB 605 AGG---ACGTCACTTCAAGAAAGCAATGTTGATGAGACTCCACTTGTGCTGTGAG 661  
QY 713 AACCAAGAAATCAATGCTGCTTTTGTGATCTGTTCTGAATGCTGCTTCAATACA 772  
DB 662 AGCCAAATGAG---TTGATCACTAGTGGCTTCTTGTCTTCACTGACGCTTCTTATA 718  
QY 773 CTACTGTCACGCTTACTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 809  
DB 719 TTACTGTCAGACTAATTTGTGTTGATGATGATGATGATGATGATGATGATGATGAT 755

Search completed: March 2, 2005, 18:17:37  
Job time : 597 sec



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 2, 2005, 11:05:37 ; Search time 185 Seconds  
(without alignments)  
7217.307 Million cell updates/sec

Title: US-10-720-018-1  
Perfect score: 816  
Sequence: 1 atgactgactacatccac.....atggtgttctactcttgg 816

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues  
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Issued Patents NA:  
2: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
7: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
8: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
9: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
10: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
11: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
12: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
13: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
14: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
15: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
16: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
17: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
18: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
19: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
20: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
21: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
22: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
23: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
24: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
25: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
26: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
27: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	238.2	29.2	696	2 US-08-336-198C-6	Sequence 6, Appli
2	238.2	29.2	696	4 US-09-184-965-6	Sequence 6, Appli
3	218.6	26.8	846	4 US-09-248-796A-29	Sequence 29, Appli
4	86.6	10.6	927	4 US-09-248-796A-2490	Sequence 2490, Appli
5	74.6	9.1	852	4 US-09-734-37B-71	Sequence 71, Appli
6	61.8	7.6	873	4 US-09-248-796A-2654	Sequence 2654, Ap
7	55	6.7	336	4 US-09-248-796A-2488	Sequence 2488, Ap
8	53	6.5	372	4 US-09-248-796A-6987	Sequence 6987, Ap
9	53	6.5	774	3 US-09-134-001C-725	Sequence 725, App
10	51.8	6.3	825	4 US-09-248-796A-2730	Sequence 2730, Ap
11	51.2	6.3	852	3 US-09-367-012-2	Sequence 2, Appli
12	51.2	6.3	852	3 US-09-777-157A-2	Sequence 2, Appli
13	51.2	6.3	852	4 US-09-734-37B-69	Sequence 69, Appli
14	51.2	6.3	951	4 US-09-489-039A-5819	Sequence 5819, Ap
15	51.2	6.3	1303	3 US-09-367-012-8	Sequence 8, Appli
16	51.2	6.3	1303	3 US-09-777-157A-8	Sequence 8, Appli
17	50.4	6.2	4411529	3 US-09-103-840A-1	Sequence 1, Appli
18	49.8	6.1	873	4 US-09-248-796A-2489	Sequence 2489, Ap
19	48.2	5.9	3993	4 US-09-710-279-3985	Sequence 3985, Ap
20	47.4	5.8	4403765	3 US-09-103-840A-2	Sequence 2, Appli
21	47.4	5.8	792	3 US-09-134-001C-1675	Sequence 1675, Ap
22	45.2	5.5	2774	3 US-09-363-189B-5	Sequence 5, Appli
23	44.4	5.4	1248	4 US-10-024-806-7	Sequence 7, Appli
24	43	5.3	879	3 US-09-468-738A-1	Sequence 1, Appli
25	43	5.3	879	3 US-09-940-019-1	Sequence 1, Appli
26	43	5.3	879	4 US-09-940-037A-1	Sequence 1, Appli
27	43	5.3	891	3 US-09-468-738A-22	Sequence 22, Appli

28	43	5.3	891	3 US-09-940-019-22	Sequence 22, Appli
29	43	5.3	891	4 US-09-940-037A-22	Sequence 22, Appli
30	43	5.3	1787	3 US-09-468-738A-17	Sequence 17, Appli
31	43	5.3	1787	3 US-09-940-019-17	Sequence 17, Appli
32	43	5.3	1787	3 US-09-940-019-17	Sequence 17, Appli
33	42.2	5.2	789	4 US-09-710-279-3157	Sequence 3157, Ap
34	42.2	5.2	3210	4 US-09-710-279-3922	Sequence 3922, Ap
35	42.2	5.2	3368	4 US-09-710-279-4169	Sequence 4169, Ap
36	41.4	5.1	825	3 US-09-134-001C-1594	Sequence 1594, Ap
37	41.4	5.1	916	4 US-09-710-279-381	Sequence 381, App
38	41.4	5.1	3024	4 US-09-710-279-3426	Sequence 3426, Ap
39	41.2	5.0	2472	4 US-09-248-796A-148	Sequence 148, App
40	41	5.0	437	3 US-09-468-738A-14	Sequence 14, Appli
41	41	5.0	437	3 US-09-940-019-14	Sequence 14, Appli
42	41	5.0	437	4 US-09-940-037A-14	Sequence 14, Appli
43	41	5.0	837	4 US-09-328-352-2671	Sequence 2671, Ap
44	40.8	5.0	601	4 US-09-949-016-62984	Sequence 62984, A
45	40.8	5.0	50263	4 US-09-949-016-13563	Sequence 13563, A

## ALIGNMENTS

RESULT 1  
US-08-336-198C-6  
Sequence 6, Application US/08336198C  
Patent No. 5866382  
GENERAL INFORMATION:  
APPLICANT: Hallborn, Johan  
APPLICANT: Penttila, Merja  
APPLICANT: Ojamo, Heikki  
APPLICANT: Keranen, Sirkka  
APPLICANT: Hann-Hagerdal, Barbei  
APPLICANT: Waldfriedson, Mats  
APPLICANT: Aitakainen, Ulla  
TITLE OF INVENTION: XYLOSE UTILIZATION BY RECOMBINANT YEASTS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasech & Birch  
STREET: 301 N. Washington St.  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-3487  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,198C  
FILING DATE: 03-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
TELEX: 248345  
INFORMATION FOR SEQ. ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 696 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULAR TYPE: cDNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Pichia stipitis  
STRAIN: CBS-6054  
FEATURE:  
NAME/KEY: CDS

LOCATION: 1..693  
 OTHER INFORMATION: /standard\_name="xy11tol  
 OTHER INFORMATION: dehydrogenase"  
 PUBLICATION INFORMATION:  
 DOCUMENT NUMBER: FI 901771  
 FILING DATE: 06-APR-1990  
 US-08-336-198C-6

Query Match 29.2%; Score 238.2; DB 2; Length 696;  
 Best Local Similarity 64.4%; Pred. No. 9.4e-65;  
 Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 227 CCTGTGATATTTCTGATTTCTGATACCGTTACAGAGTGTGTTCTCAAGTGTCTAAGATT 286  
 DB 107 CCTGCAACATTTGGGGATGCTGAGGCGATGAGCGCTTCTTACAGCTTCATCAAGAACACC 166  
 QY 287 TTGTAAAGTGTGCATTTGCACTTGTGTTAAACAGCTGTGTATCTGTGAAAACCTTCCCATGTG 346  
 DB 167 ACGGCAAGATCGGTGACTTGTGATTAAACCGCTGATCTGTGAAAACCTTCCCTGCGG 226  
 QY 347 AAGATTACCCAGCCCAAGAACGCTGAGAGATGTTAAGTTAACTTGTGGGTTCTTTGT 406  
 DB 227 AAACGTAACCCGCTACTTAACGCTGAAGCATGATGAAGTGAACGGTTGGGCTCATTTCT 286  
 QY 407 ATGTTTCTCAAGCCTTTGCTTAAGCCATTGATCAAGAGATTCAGAGGTGCTTCTGTG 466  
 DB 287 ACGTTTGGCAATCGTTCGCTAGCCATGATCCAGAA---CAACTGAGAGGCTCTATCA 343  
 QY 467 TTTTGAATTTGTTCTATGTTCTGTGTCATTTGTCAACGATCTCTCAAAACCAAGTTGTCTAA 526  
 DB 344 TCTGTATTTGCTCAAAAGTCTGGAACAAATGTCAACGACCAACCCCAATGTATGTACA 403  
 QY 527 ACATGTCCAAGGCTGTGTTATTCATTTGGCTTAAGCTTTGGCTTGTGAATGGGCTAACT 586  
 DB 404 ACATGTCCAAGGCTGTGATGATTCATTTGTGATGATCTTGGCTTGTGAATGGGCTAACT 463  
 QY 587 ACAACATCAAGATTAATTTCTTAAACCAAGTTACATCTACGCTCTTGTACCAAGATG 646  
 DB 464 ACAACATCAAGATTAATTTCTTAAACCAAGTTACATCTACGCTCTTGTACCAAGATG 523  
 QY 647 TTAATCAATGTGAACGAAGATTTGTACAAAGATGATCTCTGTATTCACCAACAAAGAA 706  
 DB 524 TGAATTTCTGGCCACACAGATGATGAAGAGCCCTGGGAATCAAGATCCCATGAAGAGAA 583  
 QY 707 TGTCCGAACCAAGAAATCATTTGTGCTGTTTGTGATCTTGTCTGATCTGTGCTT 766  
 DB 584 TGGCCGAACCCCAAGAAATCTGTGGGCTCATCTTAATCTTGGCAAGCGAGACTGCTTCTT 643  
 QY 767 CATACATCACTGTGCTGAGCTTACTGTGATGTGTGTT 805  
 DB 644 CCTACACTACGGGCCACAAATTTGTGTGTGACGGAAGAT 682

RESULT 2  
 US-09-184-965-6  
 Sequence 6, Application US/09184965  
 Patent No. 6582944  
 GENERAL INFORMATION:  
 APPLICANT: Hallborn, Johan  
 APPLICANT: Penttila, Merja  
 APPLICANT: Ojamo, Heikki  
 APPLICANT: Keranen, Sirkka  
 APPLICANT: Hahn-Hagerdal, Barbel  
 APPLICANT: Mäldfridsson, Mats  
 APPLICANT: Aitakainen, Ulla  
 TITLE OF INVENTION: XYLOSE UTILIZATION BY RECOMBINANT YEASTS  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Birch, Stewart, Kolasch & Birch  
 STREET: 301 N. Washington St.  
 CITY: Falls Church  
 STATE: Virginia  
 COUNTRY: USA

ZIP: 22046-3487  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/184,965  
 FILING DATE:  
 CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/336,198

FILING DATE: 03-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-205-8000

TELEFAX: 703-205-8050

TELEX: 248345

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 696 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Pichia stipitis

STRAIN: CBS-6054

FEATURE:

NAME/KEY: CDS

LOCATION: 1..693

OTHER INFORMATION: /standard\_name="xy11tol

PUBLICATION INFORMATION:

DOCUMENT NUMBER: FI 901771

FILING DATE: 06-APR-1990

US-09-184-965-6

Query Match 29.2%; Score 238.2; DB 4; Length 696;  
 Best Local Similarity 64.4%; Pred. No. 9.4e-65;  
 Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 227 CCTGTGATATTTCTGATTTCTGATACCGTTCAAGAGTGTGTTGCTCAAGTGTCTAAGATT 286  
 DB 107 CCTGCAACATTTGGGGATGCTGAGCGATGAGCGCTTCTTACAGCTTCATCAAGAACACC 166  
 QY 287 TTGTAAAGTGTGCATTTGCACTTGTGTTAAACAGCTGTGTATCTGTGAAAACCTTCCCATGTG 346  
 DB 167 ACGGCAAGATCGGTGACTTGTGATTAAACCGCTGATCTGTGAAAACCTTCCCTGCGG 226  
 QY 347 AAGATTACCCAGCCCAAGAACGCTGAGAGATGTTAAGTTAAGTTAACTTGTGGGTTCTTTGT 406  
 DB 227 AAACGTAACCCGCTACTTAACGCTGAAGCATGATGAAGTGAACGGTTTGGGCTCATTTCT 286  
 QY 407 ATGTTTCTCAAGCCTTTGCTTAAGCCATTGATCAAGAGATTCAGAGGTGCTTCTGTG 466  
 DB 287 ACGTTTGGCAATCGTTCGCTAGCCATGATCCAGAA---CAACTGAGAGGCTCTATCA 343  
 QY 467 TTTTGAATTTGTTCTATGTTCTGTGTCATTTGTCAACGATCTCTCAAAACCAAGTTGTCTAA 526  
 DB 344 TCTGTATTTGCTCAAAAGTCTGGAACAAATGTCAACGACCAACCCCAATGTATGTACA 403  
 QY 527 ACATGTCCAAGGCTGTGTTATTCATTTGGCTTAAGACTTGTGGCTTGTGAATGGGCTAAGT 586  
 DB 404 ACATGTCCAAGGCTGTGATGATTCATTTGGCTTGAAGCTGTGCGCTGCGAATGGGCTAAGT 463  
 QY 587 ACAATATGAGTTAATTTCTTAAACCAAGTTACATCTAGGTCCTTGTGCAAGAAATG 646  
 DB 464 ACAATATGAGTTAATTTCTTAAACCAAGTTACATCTAGGTCCTTGTGCAAGAAATG 523

QY 647 TTATCAATGTAACGAAGATTGTACACAGATGATCTGTGATCCCAACAAAGAA 706  
 DB 524 TGAATTTCTGCGCACACAGAGATGAAGAGCTGGAGATCCCAATCCCAATGAAGAA 583  
 QY 707 TGTCCGAACCAAGAAATACATGTGCTGTTTGTACTGTCTTTGTAATCTGCTT 766  
 DB 584 TGGCCGACCAAGAAATCTGTGGGTTCATCTTATCTTGGCAAGAGAGCTTTCTT 643  
 QY 767 CATACCTACTGTGTCAGCTTACTGTGTGATGTGTGTT 805  
 DB 644 CCTACCTAGCGCCCAATTTGTTGTGACGAGAT 682

RESULT 3

US-09-248-796A-29  
 ; Sequence 29, Application US/09248796A  
 ; Patent No. 6747137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keith Weinstock et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 ; FILE REFERENCE: 107196.132  
 ; CURRENT APPLICATION NUMBER: US/09/248,796A  
 ; PRIOR FILING DATE: 1999-02-12  
 ; PRIOR APPLICATION NUMBER: US 60/074,725  
 ; PRIOR FILING DATE: 1998-02-13  
 ; PRIOR APPLICATION NUMBER: US 60/096,409  
 ; NUMBER OF SEQ ID NOS: 28208  
 ; SEQ ID NO 29  
 ; LENGTH: 846  
 ; TYPE: DNA  
 ; ORGANISM: Candida albicans  
 US-09-248-796A-29

Query Match 26.8%; Score 218.6; DB 4; Length 846;  
 Best Local Similarity 62.0%; Pred. No. 1.6e-58;  
 Matches 385; Conservative 0; Mismatches 224; Indels 12; Gaps 2;

QY 14 TTCCAACTTTAGATTGATGCGCACTTAACATGTGCAAGGTGCTGTGTGTTAG 73  
 DB 218 TCCCAAGCTTCGTTGGATGAGAACTAGTATATTAACCGGTGCTGTGGTGTG 277  
 QY 74 CTGAAGCTTAACTCAAGGTTTGTGCTCAAGTTCGATCTTGTGTTGATATG 133  
 DB 278 CTGCCGTGTATCAAGGCTTTATTAAGCAAGGTCCATGTGATTAAGTCGATATGA 337  
 QY 134 ACCAAGAAAGACTGTGCAACAGCCGATATCAACAATACGCTACGTAAGAAATGA 193  
 DB 338 ACTTGAAAGAAACAAACAAGCTGCTAGAGAGCTTTACATAGGGCCGAAGACAAATGA 397  
 QY 194 AGTTGAAGAGATTCCAA-----AGATGGGTCATATGCTGTGATTTCTGATT 244  
 DB 398 AAGGTAAATACGAATACCAATCGTCAAGTGAAGTCTGTGATTAATATGGGAG 457  
 QY 245 CTGATACCTTCAAGGTGTTGCTCAAGTTCGTAAGATTTTGTAAAGTTCCTTGC 304  
 DB 458 CTGAAGCTTCGACTTGACATTCMAAGCATCAACGAACCAACGCAAAATCTCAAGT 517  
 QY 305 ACTTGTTAAACAGCTGTGTTACTGTGAAAACTTCCATGTGAAAGTTACCCAGCAAG 364  
 DB 518 TCTTGTCAACACTGCGGTGACGTGAAACTTCCAGTGAAGAGTACCCAGCAAG 577  
 QY 365 AGCTGAGAAGATGTGAAGTTAACTTGTGGGTTCTTGTATGTTTCTCAAGCTTTG 424  
 DB 578 AGCTGAAACCTTATGAAGTTAAAGGGTGGGTCATTCTACGTTTCCAAAGCTTTTG 637  
 QY 425 CTAAAGCATTTGATCAAGAAAGTATCAAGGTCCTCTGTGTTGTTGATTTGTTATGT 484  
 DB 638 CTAAAGCATTTAATCCAAA---CAACATGACCGATGATCATTTTGAATCGGGTCAATGT 694  
 QY 485 CTGATGCAATTTGATCAAGATCTCTCAAAACCAAGTTGTCTTAAACATGTCCAAGGCTGTG 544

DB 695 CCGGTACATGCTGCAACGACCAACACAAATGATGATGAACATGTCCAAAGCCGGT 754  
 QY 545 TTATCATTTGGCTAAGACTTTGGCTTGGATGGGCTTAAGTCAACATCAAGTTAAT 604  
 DB 755 TCAATCATTTAACCAGATCATTTGGCTGTGAATGGGCTAAATACATATCAGATCA 814  
 QY 605 CTTTAAACCCAGTTATACATCT 625  
 DB 815 CATTTGTCGCCCGGACATCT 835

RESULT 4

US-09-248-796A-2490  
 ; Sequence 2490, Application US/09248796A  
 ; Patent No. 6747137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keith Weinstock et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 ; FILE REFERENCE: 107196.132  
 ; CURRENT APPLICATION NUMBER: US/09/248,796A  
 ; PRIOR FILING DATE: 1999-02-12  
 ; PRIOR APPLICATION NUMBER: US 60/074,725  
 ; PRIOR FILING DATE: 1998-02-13  
 ; PRIOR APPLICATION NUMBER: US 60/096,409  
 ; NUMBER OF SEQ ID NOS: 28208  
 ; SEQ ID NO 2490  
 ; LENGTH: 927  
 ; TYPE: DNA  
 ; ORGANISM: Candida albicans  
 US-09-248-796A-2490

Query Match 10.6%; Score 86.6; DB 4; Length 927;  
 Best Local Similarity 52.0%; Pred. No. 8.1e-17;  
 Matches 328; Conservative 0; Mismatches 279; Indels 24; Gaps 5;

QY 182 CTGAAGATTTGAAGTTGAAAGATTCACAAAGTGGTTCAATGCTGTGATTTCTG 241  
 DB 308 CTGAATATTTAAACAAGACTTATGGGTTAAATCCAAAGCATCAAAATGTTAATCTG 367  
 QY 242 ATTCTGATACCTTCAACAAGGTTTGTCTCAAGTTGCTAAGATTTGTAAGTTCAT 301  
 DB 368 ATTTCCAGATTTGAAAGTTTGTCAACAATTTGAATGATGATTTGCTATTTGATGATA 427  
 QY 302 TGCATGTTGTTAAACAGCTGTGTTCTG--TGAAACTTCCATGTGAAGTTACCCAG 358  
 DB 428 TCTTTGTTCCATGCTGTGTTGCTGGAACGAAAGGCCGAAATGATCAAGGAG 487  
 QY 359 CCAAGAACCTGAGAAGATGTGAAGTTAACTTGTGGTCTTTGTATGTTTCTCAAG 418  
 DB 488 TCGACAAATGAAACAAGTTGTTGATTTGAATTTAAACAGTGTATTAATTTGTCTCATG 547  
 QY 419 CCTTGTGTAAGCATTTGATCAAGAAGTATCAAGGTCCTGTGTTTGAATGTTGTT 478  
 DB 548 TTGTTGTTCAATTTTCAGAAAAGGTTA-----AAGGTCATTTCAATTTCACTGCA 601  
 QY 479 CTATGCTGTGTCATTTGTCAGATCTCTCAAAACCAAGTTGTTCTCAACATGTCCAG 538  
 DB 602 GTATGTCGCTTCAATTTGTTAATGTCACAAATTTCAAGCACTTAAACAGCTGTAAG 661  
 QY 539 CTGCTGTATTCATTTGGGTTAAGACTTTGGCTTGTGAATGGCTAAGTCAACATCAAG 598  
 DB 662 CTGGGGTCAAAATTTGTCAAATCATTAAGTGTGAAGGGCAACATTTGCT---AGAG 718  
 QY 599 TTAATTTCTTAAACCAAGTTACATCTAGCTCTTGTGCAAGAATGTTATCAATGTA 658  
 DB 719 TCAATTTCTTTCTCAAGTTACATGCTACTCATCTTAAGTAATTTGCTG----- 769  
 QY 659 AGAAGATTTGATCAACAGATGATCTGTGATCCCAACAAAGATGTCCAGCAA 718  
 DB 770 ATCCGATGTCAAGAGTAAATGTTGCAACTTAACACACTGTGTAGAAAGCCAAACCA 829





NUMBER OF SEQ ID NOS: 28208  
 ; SEQ ID NO 2488  
 ; LENGTH: 336  
 ; TYPE: DNA  
 ; ORGANISM: Candida albicans  
 US-09-248-796A-2488

Query Match 6.7%; Score 55; DB 4; Length 336;  
 Best Local Similarity 52.9%; Pred. No. 4.3e-07;  
 Matches 171; Conservative 0; Mismatches 140; Indels 12; Gaps 2;

QY 491 CCATTGTCAACGATCTCTCAAAACCAAGTGTCTACAAACATGTCACAGCTGTGTATCC 550  
 DB 20 CCCGCTGTGAATACCTCAATTCACAGCTCTTATATAGCTGTAAAGCTGCATGATC 79  
 QY 551 ATTGGCTAAGACTTGGCTGTGATGAGTGGCTTAAGTACCAATCAGAGTTAATCTTTAA 610  
 DB 80 ATTAGCCAAATCATTTGAGTGTGGAATGGCTAGTTTGTGTCTAGATTAATTAATTT 139  
 QY 611 ACCCAGGTACATCTACGCTCTTTTACCAAGAAATGTATCAATGTGTACGAAGATTTGT 670  
 DB 140 CTCACAGGATATTA-----TTGACTGATATTTCTGATTTTGTGATCCGAATATA 190  
 QY 671 ACAACAGATGATCTCTGATCCCAACAAGAAATGTCCGAACCAAGAAATATCATTTG 730  
 DB 191 AAAAGAAATGTGTGCAATGACACCTTTGGAGAGAGAGATTACACAGAAATTTGTGG 250  
 QY 731 GTGCTGTTTGTACTGCTTTCTGAATCTGCTGCTTCATACACTAGTGTGCCAGCTTAC 790  
 DB 251 GGGCATATTTTAACTTGGCTC---AAATGCATCACTTAATACTAGTGTCCAAATATTG 307  
 QY 791 TGTGTATGTGTGTTCACTTCT 813  
 DB 308 CTGTGATGGGGTTATATACATGT 330

RESULT 8

US-09-248-796A-6987/c  
 ; Sequence 6987, Application US/09248796A  
 ; Patent No. 6747137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keith Weinstock et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.132  
 ; CURRENT APPLICATION NUMBER: US/09/248,796A  
 ; PRIOR FILING DATE: 1999-02-12  
 ; PRIOR APPLICATION NUMBER: US 60/074,725  
 ; PRIOR FILING DATE: 1998-02-13  
 ; PRIOR APPLICATION NUMBER: US 60/096,409  
 ; NUMBER OF SEQ ID NOS: 28208  
 ; SEQ ID NO 6987  
 ; LENGTH: 372  
 ; TYPE: DNA  
 ; ORGANISM: Candida albicans  
 US-09-248-796A-6987

Query Match 6.5%; Score 53; DB 4; Length 372;  
 Best Local Similarity 62.4%; Pred. No. 2e-06;  
 Matches 83; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 14 TTCCAACTTTAGATTGATGCGCACTTAACATTTGTCAAGGTGCTGTGTGTTTAA 73  
 DB 205 TCCCAAGCTTCCGTTTGATGAGAAACTAGTCATATTAACCGGTGCTGTGTGTTGG 146  
 QY 74 CTGAAGCTTAAATCAAGGGTGTGGCTTGAAGTCTGACATTCATTTGCTTGAATG 133  
 DB 145 CTGCCGTGTATCAAGAGCTTATTAAGCCAAAGTCCGATGTGATTAAGTATGA 86  
 QY 134 ACCAAGAAAAAGAC 146  
 DB 85 ACTTGAAAAAGAC 73

RESULT 9  
 US-09-134-001C-725  
 ; Sequence 725, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; PRIOR FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 725  
 ; LENGTH: 774  
 ; TYPE: DNA  
 ; ORGANISM: Staphylococcus epidermidis  
 US-09-134-001C-725

Query Match 6.5%; Score 53; DB 3; Length 774;  
 Best Local Similarity 47.8%; Pred. No. 3e-06;  
 Matches 223; Conservative 0; Mismatches 235; Indels 9; Gaps 2;

QY 200 AAGAAATTCCAAGATGGGTTCATATGCTGTGATATTTCTGATTCGATACCGTTTACA 259  
 DB 182 AATCAGGTGTTAAACACATCATTAATTAATGATTAATCATCATGATGATGAATATC 241  
 QY 260 AGGTGTTGCTCAAGTGTCTTAAGATTTTGTGAAGTGTGCACTTGTGTTACACAG 319  
 DB 242 AATGTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 298  
 QY 320 CTGGTTACTGTGATAAATCTCCATGGAAGATTAACCCAGCAAGAACGTCGAAGATG 379  
 DB 299 CTAGATATCTCAATTCAGATGATGATGATGATGATGATGATGATGATGATGATG 358  
 QY 380 TGAAGTTAACTGTGTTGCTTTGATGATGTTCTCAAGCTTTGCTGAAGCATGATCA 439  
 DB 359 TCAATTAAGCTTAATATGATGATGATGATGATGATGATGATGATGATGATGATG 418  
 QY 440 AAGAAATGTAAGAGGTCTTCTGTTGTTTGAATGTTCTATGCTGTGTCATTTGCA 499  
 DB 419 AAGAGGCA-----GTGGTTGATGATGATGATGATGATGATGATGATGATGATG 472  
 QY 500 AGATTCCTAAACCAAGTGTCTCAACATGTCAGAGCTGTGATTCATTTGGCTA 559  
 DB 473 ATTAACGCAAGACCAAGCTCTTATGAACGTTAAAGCCGTGTGACAGATGATGAT 532  
 QY 560 AGACTTGGCTGTGAATGGCTTAAGTACACATGACAGTTAATCTTTAAACCGAGTT 619  
 DB 533 AAGCTTACTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 592  
 QY 620 ACATCTACGCTCTTTGACCAAGAAATGTTATCAATGTTACCAAGAA 666  
 DB 593 ACATGGAACAATCGAAGCAAGAAATGATTAATGATTAATGATGATGATGATGAT 639

RESULT 10

US-09-248-796A-2730  
 ; Sequence 2730, Application US/09248796A  
 ; Patent No. 6747137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keith Weinstock et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.132  
 ; CURRENT APPLICATION NUMBER: US/09/248,796A  
 ; PRIOR FILING DATE: 1999-02-12  
 ; PRIOR APPLICATION NUMBER: US 60/074,725





his Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 2, 2005, 04:46:41 ; Search time 560 Seconds  
(without alignments) 8625.909 Million cell updates/sec

Title: US-10-720-018-1

Perfect score: 816  
Sequence: 1 atgactgactacatcccaac.....atggtggttcactctcttgg 816

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1808:\*  
2: geneseqn19808:\*  
3: geneseqn20008:\*  
4: geneseqn20018:\*  
5: geneseqn20018:\*  
6: geneseqn20028:\*  
7: geneseqn20028:\*  
8: geneseqn20038:\*  
9: geneseqn20038:\*  
10: geneseqn20038:\*  
11: geneseqn20038:\*  
12: geneseqn20048:\*  
13: geneseqn20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81.6	100.0	816	12 ADQ94601	Adg94601 Ambrosioz
2	99.4	12.2	3303	3 AAF07518	Aaf07518 Fusarium
3	88.6	10.9	687	3 AAF12409	Aaf12409 Aspergill
4	81.8	10.0	637	13 ADS61049	Ads61049 Bacterial
5	80.6	9.9	735	13 ADS48561	Ads48561 Bacterial
6	74.6	9.1	852	5 AAH74598	Aah74598 Synthetic
7	74.2	9.1	1146	3 AAF12302	Aaf12302 Aspergill
8	67.6	8.3	783	13 ADS47380	Ads47380 Bacterial
9	66.4	8.1	789	12 ADH62512	Adh62512 A. thalia
10	63.6	7.8	774	13 ADS48180	Ads48180 Bacterial
11	60.6	7.4	636	3 AAF11343	Aaf11343 Aspergill
12	59.8	7.3	822	9 AAL57425	Aal57425 Datura st
13	59	7.2	774	3 AAZ46762	Aaz46762 Bacillus
14	59	7.2	2255	3 AAZ46763	Aaz46763 Bacillus
15	56.4	6.9	798	3 AAC42189	Aac42189 Arabidops
16	56.4	6.9	825	9 AAL57426	Aal57426 Hyoscyam
17	56.2	6.9	789	3 AAC43167	Aac43167 Arabidops
18	56.2	6.9	789	6 ABZ13143	Abz13143 Arabidops
19	56.2	6.9	789	12 ADH62514	Adh62514 A. thalia
20	55.2	6.8	732	13 ADS63286	Ads63286 Bacterial

21	55.2	6.8	732	13 ADS63285	Ads63285 Bacterial
22	55.2	6.8	943	12 ADN62526	Adn62526 A. thalia
23	55	6.7	949	6 ABN98571	Abn98571 Arabidops
24	54.4	6.7	783	3 AAC43166	Aac43166 Arabidops
25	54.4	6.7	783	6 ABZ13220	Abz13220 Arabidops
26	53.8	6.6	768	13 ADS48201	Ads48201 Bacterial
27	53.8	6.6	1003	12 ADN62520	Adn62520 A. thalia
28	53.2	6.5	852	12 ADJ98111	Adj98111 Ketoreduc
29	53.2	6.5	852	12 ADJ98097	Adj98097 Ketoreduc
30	53.2	6.5	855	12 ADJ98103	Adj98103 Ketoreduc
31	53.2	6.5	855	12 ADJ98105	Adj98105 Ketoreduc
32	53	6.5	774	6 ABN91262	Abn91262 Staphyloc
33	53	6.5	774	13 ADS01477	Ads01477 Staphyloc
34	52.8	6.5	819	11 ADJ12063	Adj12063 Maize cDN
35	52.8	6.5	1366	12 ADN62538	Adn62538 A. thalia
36	52.4	6.4	656	6 ABQ65920	Abq65920 Arabidops
37	52.4	6.4	789	12 ADN62500	Adn62500 A. thalia
38	52.2	6.4	852	12 ADJ98109	Adj98109 Ketoreduc
39	51.6	6.3	852	12 ADJ98115	Adj98115 Ketoreduc
40	51.6	6.3	852	12 ADJ98113	Adj98113 Ketoreduc
41	51.6	6.3	855	12 ADJ98101	Adj98101 Ketoreduc
42	51.4	6.3	849	6 ABZ32459	Abz32459 Candida a
43	51.2	6.3	711	11 ADJ12134	Adj12134 Maize cDN
44	51.2	6.3	810	6 ABZ13885	Abz13885 Arabidops
45	51.2	6.3	810	8 ABZ42061	Abz42061 Arabidops

## ALIGNMENTS

RESULT 1	ADQ94601	standard; cDNA; 816 BP.
ID	ADQ94601	
AC	ADQ94601	
XX		
DT	23-SEP-2004	(first entry)
XX		
DE	Ambrosiozyma monospora	NADH dependent L-xylose reductase cDNA.
XX		
KW	NADH dependent L-xylose reductase; fermentation; biomass; ALX1; gene;	
XX		
OS	Ambrosiozyma monospora.	
XX		
XX		
FT	CDS	Location/Qualifiers
FT		1..816
FT		/*tag= a
FT		/product= "Ambrosiozyma monospora NADH dependent L-
FT		xylose reductase"
FT		/partial
FT		/note= "No stop codon"
FT		/EC_number= "EC 1.1.1.10"
XX		
XX		
XX	US2004132074-A1.	
XX		
XX	08-JUL-2004.	
XX		
XX	24-NOV-2003; 2003US-00720018.	
XX		
PR	16-FEB-2001; 2001FI-00000308.	
PR	15-FEB-2002; 2002MO-FI000125.	
PR	10-MAR-2003; 2003US-00257821.	
PR	12-SEP-2003; 2003FI-00001307.	
XX		
PA	(VALW ) VALTION TEKNIILINEN TUTKIMUSKESKUS.	
XX		
PI	Verno R, Richard P, Penttila M;	
XX		
XX	WPI; 2004-517001/49.	
DR	P-PSDB; ADQ94602.	
XX		
PT	New DNA molecule encoding NADH dependent L-xylose reductase, useful in	

PT producing fermentation products or in conversion of cheap biomass to  
PT useful products.  
XX  
PS Claim 6; SEQ ID NO 1; 15pp; English.  
XX  
CC The present invention relates to a NADH dependent L-xylulose reductase  
CC enzyme protein and its encoding polynucleotide. The invention is useful  
CC in producing fermentation products or in conversion of cheap biomass to  
CC useful products. The present sequence is an Ambrosiozyma monospora NRRL X  
CC -1484 NADH dependent L-xylulose reductase cDNA, A1X1 (EC 1.1.1.10).  
XX  
SQ Sequence 816 BP; 217 A; 159 C; 182 G; 258 T; 0 U; 0 Other;  
Query Match 100.0%; Score 816; DB 12; Length 816;  
Best Local Similarity 100.0%; Pred. No. 2.7e-227;  
Matches 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGACTGACTACATTCCTTAACTTGAATTCGATGCGCACTTAAACCATGTCACAGGTGCC 60  
DB 1 ATGACTGACTACATTCCTTAACTTGAATTCGATGCGCACTTAAACCATGTCACAGGTGCC 60  
QY 61 TGTGGGTGTTAGCTGCAAGCTTAACTTGAATTCGATGCGCACTTAAACCATGTCACAGGTGCC 120  
DB 61 TGTGGGTGTTAGCTGCAAGCTTAACTTGAATTCGATGCGCACTTAAACCATGTCACAGGTGCC 120  
QY 121 TTGCTTGATATCGACCAAGAAAAGACTGTCGCAAAACAAGCAATATACCAATATACGCT 180  
DB 121 TTGCTTGATATCGACCAAGAAAAGACTGTCGCAAAACAAGCAATATACCAATATACGCT 180  
QY 181 ACTGAAGAAATGAAGTTGAAGAAAGTTCCAAAGATGGGTTCAATATGCTGTGAATATTTCT 240  
DB 181 ACTGAAGAAATGAAGTTGAAGAAAGTTCCAAAGATGGGTTCAATATGCTGTGAATATTTCT 240  
QY 241 GATTCGATACCGTTCAAGAGGTTTGTCAAGTTGCTAAGATTTTGTGTAAGTTGCCA 300  
DB 241 GATTCGATACCGTTCAAGAGGTTTGTCAAGTTGCTAAGATTTTGTGTAAGTTGCCA 300  
QY 301 TTGCACTTGGTTAAACAACAGCTGGTTACTGTGAAAACCTTCCATGTAAGATTAACCGCC 360  
DB 301 TTGCACTTGGTTAAACAACAGCTGGTTACTGTGAAAACCTTCCATGTAAGATTAACCGCC 360  
QY 361 AAGAAACGCTGAAGAAAGTGTGAAGTTAACTTGGTGGTTCTTGTATGTTCTCAAGCC 420  
DB 361 AAGAAACGCTGAAGAAAGTGTGAAGTTAACTTGGTGGTTCTTGTATGTTCTCAAGCC 420  
QY 421 TTGCTAAGCCATGATCAAGAGGTATCAAGGGCTTCTGTTGTTTGAATGGTCT 480  
DB 421 TTGCTAAGCCATGATCAAGAGGTATCAAGGGCTTCTGTTGTTTGAATGGTCT 480  
QY 481 ATGTCTGTGTCATTCATCAAGATCCTCAAAACAAGTTGTCACAACTGTCCAAAGCT 540  
DB 481 ATGTCTGTGTCATTCATCAAGATCCTCAAAACAAGTTGTCACAACTGTCCAAAGCT 540  
QY 541 GGTGTATTCATTTGGCTTAAGACTTTGGCTTGAATGGGCTAAGTACCAATCAAGATT 600  
DB 541 GGTGTATTCATTTGGCTTAAGACTTTGGCTTGAATGGGCTAAGTACCAATCAAGATT 600  
QY 601 AATTCTTTAAACCAAGTTACATCAAGCTTCTTGAACCAAGAATGTAATCAATGTAAC 660  
DB 601 AATTCTTTAAACCAAGTTACATCAAGCTTCTTGAACCAAGAATGTAATCAATGTAAC 660  
QY 661 GAAGAATTGTACAACAGATGATCTCTGATATCCCAACAAGAATGTCGAACCAAG 720  
DB 661 GAAGAATTGTACAACAGATGATCTCTGATATCCCAACAAGAATGTCGAACCAAG 720  
QY 721 GAATACATTTGGCTGTTTGTACTGCTTGTGAATTCGCTGCTTCACTACCTAGT 780  
DB 721 GAATACATTTGGCTGTTTGTACTGCTTGTGAATTCGCTGCTTCACTACCTAGT 780  
QY 781 GCCAGCTTACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 816  
DB 781 GCCAGCTTACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 816

RESULT 2  
AAAF07518  
ID AAFA07518 standard; cDNA, 3203 BP.  
XX  
XX AAFA07518;  
XX  
DT 13-MAR-2001 (first entry)  
XX  
DE Fusarium venenatum EST SEQ ID NO:41.  
XX  
XX Multiple gene expression; filamentous fungal cell; EST;  
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
KW culture condition; environmental stress; spore morphogenesis;  
KW metabolic pathway engineering; catabolic pathway engineering; ss.  
XX  
OS Fusarium venenatum.  
XX  
PN WO200056762-A2.  
XX  
PD 28-SEP-2000.  
XX  
PF 22-MAR-2000; 2000WO-US007781.  
XX  
PR 22-MAR-1999; 99US-00273623.  
XX  
PA (NOVO ) NOVO NORDISK BIOTECH INC.  
PA (NOVO ) NOVO NORDISK AS.  
XX  
PI Berka RM, Rey WM, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
XX  
DR WPI: 2000-594572/56.  
XX  
PT Monitoring differential expression of genes in filamentous fungal cells  
PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
XX substrate of expressed sequence tags.  
XX  
PS Claim 86; Page 383-384; 3161pp; English.  
XX  
CC The present invention describes a method for monitoring differential  
CC expression of genes in a first filamentous fungal (FF) cell relative to  
CC expression of the same genes in one or more second filamentous fungal  
CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
CC are used in the methods for monitoring differential expression of genes  
CC in a first filamentous fungal (FF) cell relative to expression of the  
CC same genes in one or more second filamentous fungal cells. Monitoring the  
CC global expression of genes from FF cells allows the production potential  
CC of the microorganisms to be improved. New genes may be discovered,  
CC possible functions of unknown open reading frames can be identified and  
CC gene copy number variation and stability can be monitored. The expression  
CC of genes can be used to study how FF cells adapt to changes in culture  
CC conditions, environmental stress, spore morphogenesis, recombination,  
CC metabolic or catabolic pathway engineering. Using ESTs provides several  
CC advantages over genomic or random cDNA clones including elimination of  
CC redundancy as one spot on an array equals one gene or open reading frame,  
CC and organization of the microarrays based on function of the gene  
CC products to facilitate analysis of the results. AAFA07478 to AAFA11247  
CC represents ESTs from Fusarium venenatum; AAFA11248 to AAFA11853 represents  
CC ESTs from Aspergillus niger; AAFA11854 to AAFA14878 represents ESTs from  
CC Trichoderma reesei; and AAFA14879 to AAFA15337 represents ESTs from  
CC Trichoderma reesei, which are all specifically claimed in the present  
XX invention  
XX  
SQ Sequence 3203 BP; 770 A; 861 C; 750 G; 815 T; 0 U; 7 Other;  
Query Match 12.2%; Score 99.4; DB 3; Length 3203;  
Best Local Similarity 52.7%; Pred. No. 4.5e-18;  
Matches 265; Conservative 0; Mismatches 231; Indels 7; Gaps 2;  
QY 308 TGGTTAACACAGCTGTTACTGTGAAAACCTTCCATGTGAAGATTACCAACCAAGAACG 367  
DB 308 TGGTTAACACAGCTGTTACTGTGAAAACCTTCCATGTGAAGATTACCAACCAAGAACG 367

Db 2614 TAGTCAGTCGCTGGCTTCACGAGAAATTTGAGGCAATCACTATCCCATGATCGTA 2673  
 Qy 368 CTGAGAAAGTGTGAAGTTAACTTGTGGCTTCTTGATGTTTCTCAAGCCTTTGCTA 427  
 Db 2674 TGGCGAAGTGTGGGGTGTAAATGTGACGCTACTTACCTCTTGTGAGTTGGCCCA 2733  
 Qy 428 AGCCATGATCAAAAGAAAGTATCAAGGGGCTTCTGTGTGTTTGTGATGTTGTTATGTCG 487  
 Db 2734 A--ACATCTCATGAAACGCCAGGTACCTGTGTGTTGTTGTTATTTGAAAGCATGTCG 2790  
 Qy 488 GTGCCATTTGTCAACGATTCCTCAAAACCAAGTTGTCTACAACATGTCCAAAGCTGTGTGA 547  
 Db 2791 GTGCTATTTGTCAATGTTTCCACAGCCACAGCCACATTAACGCGGCGAAGGCGCTGTC 2850  
 Qy 548 TCCATTTGGCTTAAGACTTTGGCTTGTGATGGGCTTAAGTCAACATCAGATTATTTCT 607  
 Db 2851 GTCACTGGCTGTCTCCCTCGCAGTGTAGAGGGGCTACGCTGGAATCCGATCACTGTA 2910  
 Qy 608 TAAACCCAGGTTACATCTACGCTCTTGTGACCAAGATGTTATCANTGGTAAAGAAAT 667  
 Db 2911 TCTCTCTGGCTACATGTTGACTGCTCTTAACGCAAGATTTGAAAGATTAACCCAGATA 2970  
 Qy 668 TGTACACAGATGATCTCTGTATCCCAACAAAGATGTCCGAACCAAGAAATACA 727  
 Db 2971 TCGAGAGACCTGGACATCTCTATTCTCAGGAGAGTATGGAGACTACCTCAAGATCTGA 3030  
 Qy 728 TTGTGCTGTTTGTACTGCTTCTGATCTGCTGCTTCAATCACTACTGTCGCACT 787  
 Db 3031 TGGGTCTGT---GACCTTTTGTGTCAGATGCTGCTTCTTCAATGATCGGGGAGATC 3086  
 Qy 788 TACTGTTGATGTTGCTTCACT 810  
 Db 3087 TTAGATTAATGAGATATATCT 3109  
 RESULT 3  
 AAF12409  
 ID AAF12409 standard; cDNA; 687 BP.  
 AC AAF12409;  
 DT 13-MAR-2001 (first entry)  
 XX Aspergillus oryzae EST SEQ ID NO:4932.  
 DE Multiple gene expression; filamentous fungal cell; EST;  
 KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
 KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
 KW culture condition; environmental stress; spore morphogenesis;  
 KW metabolic pathway engineering; catabolic pathway engineering; ss.  
 OS Aspergillus oryzae.  
 XX  
 PN WO200056762-A2.  
 XX  
 PD 28-SEP-2000.  
 XX  
 PF 22-MAR-2000; 2000MO-US007781.  
 XX  
 PR 22-MAR-1999; 99US-00273623.  
 XX  
 PA (NOVO ) NOVO NORDISK BIOTECH INC.  
 PA (NOVO ) NOVO NORDISK AS.  
 XX  
 PI Berka RM, Rey MW, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;  
 DR WPI; 2000-594572/56.  
 XX  
 PT Monitoring differential expression of genes in filamentous fungal cells  
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
 PT substrate of expressed sequence tags.  
 XX  
 PS Claim 88; Page 2078; 3161pp; English.

XX The present invention describes a method for monitoring differential  
 CC expression of genes in a first filamentous fungal (FP) cell relative to  
 CC expression of the same genes in one or more second filamentous fungal  
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
 CC the FP cells and a substrate of expressed sequence tags (EST). The ESTs  
 CC are used in the methods for monitoring differential expression of genes  
 CC in a first filamentous fungal (FP) cell relative to expression of the  
 CC same genes in one or more second filamentous fungal cells. Monitoring the  
 CC global expression of genes from FP cells allows the production potential  
 CC of the microorganisms to be improved. New genes may be discovered,  
 CC possible functions of unknown open reading frames can be identified and  
 CC gene copy number variation and stability can be monitored. The expression  
 CC of genes can be used to study how FP cells adapt to changes in culture  
 CC conditions, environmental stress, spore morphogenesis, recombination,  
 CC metabolic or catabolic pathway engineering. Using ESTs provides several  
 CC advantages over genomic or random cDNA clones including elimination of  
 CC redundancy as one spot on an array equals one gene or open reading frame,  
 CC and organization of the microarrays based on function of the gene  
 CC products to facilitate analysis of the results. AAF07478 to AAF11247  
 CC represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents  
 CC ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from  
 CC *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from  
 CC *Trichoderma reesei*, which are all specifically claimed in the present  
 CC invention  
 XX  
 SQ Sequence 687 BP; 152 A; 210 C; 175 G; 147 T; 0 U; 3 Other;  
 Query Match 10.9%; Score 88.6; DB 3; Length 687;  
 Best Local Similarity 49.8%; Pred. No. 3.3e-15;  
 Matches 279; Conservative 0; Mismatches 275; Indels 6; Gaps 2;  
 Qy 196 TTGAAGAAGTTCAAAGATGGGTTATGCTGTGATTTCTGATTTCTGATACCTT 255  
 Db 114 TTGAACAATGCTCCCAACCTCAACCGCCCTACGCTGATGTTCCACCTTACTCCGTC 173  
 Qy 256 CACAAGGTTTGTCCCAAGTGTCTAAGATTTGTGTAAGTTCATGCTGTTGTTAAC 315  
 Db 174 AACGATGCCCTCTCGATATTTATCTCAAGCAGCGCAAT---CGACAACCTGTGTACAC 230  
 Qy 316 ACAGCTGTACTGTGAAAACCTTCCCATGTGAATTAACCCAGCCAAAGACGTGAGAAG 375  
 Db 231 TCGCCGATTCACGGAACCTTCATGCAATCTCTTACCTCAAGACCGTGTGCAAAAG 290  
 Qy 376 ATGTGAAGTTAATCTGTTGGTCTTGTGATGTTTCTCAAGCCTTGTCTAAGCCATTG 435  
 Db 291 CTTTGGGGGCTTAATGTGATGGAACATACCTTTTCCGACCGGTGTCCCAAGC---AC 347  
 Qy 436 ATCAAGAAGTATCAAGGTCCTTGTGTTGATGTTGTTGATGTTGATGTTGATGTTGATG 495  
 Db 348 CTCATGAGCGCAAGGTTCCGACGACATGTGATGATGATGATGATGATGATGATGATG 407  
 Qy 496 GTCAACGATCTCAAAACCAAGTTGCTCAACATGTCCAGAGCTGTGTTATTCATTG 555  
 Db 408 GACAAGTCCCGACGCCAGGCTCTTACAAAGCCGCAAGCCGCTGTGTCAACTT 467  
 Qy 556 GCTAAGACTTTGCTTGTGAATGGCTTAAGTCAACATCAAGTTATTTCTTTAAACCA 615  
 Db 468 GCCGCTACTTGGCCGCGCAATGGCGCGGTCAACATCCGGGTGAGCTGACGACCCCT 527  
 Qy 616 GGTTCATCTACGGTCTTGTGACCAAGATGTTATCAATGATGATGATGATGATGATGATG 675  
 Db 528 CGATACATCTTACCTGCTGACCGCGCAAGATTTGATGATGATGATGATGATGATGATG 587  
 Qy 676 AGATGATCTGTGATCCCAACAAGATGTCCGAACCAAGAAATATATTTGTGCT 735  
 Db 588 AAGTGATCTGCTATCCCAACCGGCAAGATGGGATCCCGAGAGACCTGATGGTCTCC 647  
 Qy 736 GTTTGTACTTGTCTTCTGA 755  
 Db 648 GTTACCTTCTGCTCAGNGA 667





CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transgenic plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polynucleotide used in  
CC the scope of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX SQ Sequence 735 BP; 136 A; 234 C; 183 G; 182 T; 0 U; 0 Other;

Query Match 9.9%; Score 80.6; DB 13; Length 735;

Best Local Similarity 61.8%; Pred. No. 7.4e-13;

Matches 18; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 436 ATCAAGAAGTATCAAGGGTCTTCTGTTGTTGATGTTCTATGCTGTCATT 495  
DB 364 ATGATGCGCGCTGGCAACCGGCTCATATTCCTTGTCCGATGTCGGACATC 423  
QY 496 GTCAACGATCTCTCAAAACCAAGTTGTCTACACATGTCGAAGCTGTGTTATCCATTG 555  
DB 424 GTCAATACCTCTCAGAGAGAGTGTCTCAACGCTCTAAGGCTGTGTCATTGACGTC 483  
QY 556 GCTAAGACTTGGCTGTGAATGGGTAAGTACATGATGAGTTAATCTTAAACCA 615  
DB 484 GGCAGATCTCTTGTCTCCAGTGGGCCAAGTACATCCGTGTACTGCAATTTCCCC 543  
QY 616 GGTACATCTACCGTCTTGTGACCAAG 642  
DB 544 GGTACATGATACCGCTCTGATTAAG 570

RESULT 6

AAH74598

ID AAH74598 standard; DNA; 852 BP.

XX AC AAH74598;

DT 15-OCT-2001 (first entry)

DE Synthetic nucleotide sequence of a NADPH-dependent carbonyl reductase.

KW Methionine gamma-lyase; mda gene, free folding energy; gene shuffling;  
KW directed evolution; molecular breeding;  
KW NADPH-dependent carbonyl reductase; ss.

OS Candida magnoliae.

XX WO200155342-A2.

PN 02-AUG-2001.

PF 31-JAN-2001; 2001WO-US003186.

PR 31-JAN-2000; 2000US-00494921.  
PR 08-DEC-2000; 2000US-00734237.

(BIOC-) BIOCATALYTICS INC.

XX

PI Rozzell DJ, Bui P, Hua L;

XX WPI; 2001-483235/52.

PT Designing synthetic nucleic acid sequences for improved amplification,  
PT expression in host cell, by comparing free energy of folding of a  
PT starting polynucleotide and a modified polynucleotide having a codon  
PT replacement.

XX Claim 9; Page 104, 117pp; English.

XX The present sequence encodes a modified NADPH-dependent carbonyl  
CC reductase. The polynucleotide was modified using the method of the  
CC invention. The specification describes a method for designing a synthetic  
CC polynucleotide. The method comprises providing a starting polynucleotide,  
CC determining the predicted free energy of folding per base of the  
CC polynucleotide, modifying the polynucleotide by replacing a codon with a  
CC different codon to provide a modified polynucleotide, determining free  
CC energy of folding per base of the modified polynucleotide, and comparing  
CC this with that of the original polynucleotide. The method is useful for  
CC developing nucleic acid sequences that enhance expression of the encoded  
CC protein in a heterologous host. The design and preparation of the  
CC synthetic genes are used in application of gene shuffling, directed  
CC evolution and molecular breeding methods. The method allows expression of  
CC genes from various organisms such as mammals, plants, yeast, fungi and  
CC bacteria in prokaryotic hosts, such as *Escherichia coli* and eukaryotic  
CC hosts at commercially viable levels, in particular proteins with low  
CC yield such as methionine gamma-lyase from *P. putida*

XX SQ Sequence 852 BP; 196 A; 216 C; 205 G; 235 T; 0 U; 0 Other;

Query Match 9.1%; Score 74.6; DB 5; Length 852;

Best Local Similarity 56.0%; Pred. No. 4.4e-11;

Matches 209; Conservative 0; Mismatches 149; Indels 15; Gaps 3;

QY 437 TCAAGAAGTATCAAGGGTCTTCTGTTGTTGATGTTCTATGCTGTCATT 496  
DB 485 TCAAGAAGTATCAAGGGTCTTCTGTTGATGTTCTATGCTGTCATT 544  
QY 497 TCAAGGATCTCAAAACCAAGTTGTCTACACATGTCGAAGCTGTGTTATCCATTG 556  
DB 545 TTAAGGTACCGCAGTTTCAGGCTACCTCAACGCTGTAAAGCTGTGTCATTG 604  
QY 557 CTAAAGCTTGGCTGTGAATGGGTAAGTACATGATGAGTTAATCTTAAACCA 616  
DB 605 CTAAATCTCTGCTGTGAATGGGTAAGTACATGATGAGTTAATCTTAAACCA 661  
QY 617 GTTACATCTACCGTCTTGTGACCAAGAAATGTTATGATGTAACGAATGTAACA 676  
DB 662 GCTACATCAAC-----ACCGAAATCTGATCTTGTATACCGAGAAATCAAGACA 712  
QY 677 GATGATCTCTGATATCCCAACAAGAAATGTCGAACCAAGAAATCAATGTCGTC 736  
DB 713 AATGGGATCTCTGATATCCCGTGGCCGATGGGAAACCTGTAACGTTGGTGT 772  
QY 737 TTTTACTCTGCTTGTGAATGCTGCTTATACATCACTAGTGCACGTTACTGTTG 796  
DB 773 AACTGTCTCTGCTTGTGA---CGCTGTCTTCAAGCTACCGGCACTGACATCATGTTG 829  
QY 797 ATGGTGGTTTAC 809  
DB 830 ACGGTGTTAC 842

RESULT 7

AAFI2302

ID AAFI2302 standard; cDNA; 1146 BP.

XX AC AAFI2302;

DT 13-MAR-2001 (first entry)

DE *Aspergillus oryzae* EST SEQ ID NO:4825.

XX Multiple gene expression; filamentous fungal cell; EST;  
 KM expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
 KM Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
 KM culture condition; environmental stress; spore morphogenesis;  
 KM metabolic pathway engineering; catabolic pathway engineering; ss.  
 XX Aspergillus oryzae.  
 OS  
 PN WO20056762-A2.  
 XX  
 PD 28-SEP-2000.  
 XX  
 PF 22-MAR-2000; 2000WO-US007781.  
 XX  
 PR 22-MAR-1999; 99US-00273623.  
 XX  
 PA (NOVO ) NOVO NORDISK BIOTECH INC.  
 PA (NOVO ) NOVO NORDISK AS.  
 PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
 XX WPI; 2000-594572/56.  
 DR  
 PT Monitoring differential expression of genes in filamentous fungal cells  
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
 PT substrate of expressed sequence tags.  
 XX  
 PS Claim 88; Page 2039; 3161pp; English.  
 XX  
 CC The present invention describes a method for monitoring differential  
 CC expression of genes in a first filamentous fungal (FF) cell relative to  
 CC expression of the same genes in one or more second filamentous fungal  
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
 CC are used in the methods for monitoring differential expression of genes  
 CC in a first filamentous fungal (FF) cell relative to expression of the  
 CC same genes in one or more second filamentous fungal cells. Monitoring the  
 CC global expression of genes from FF cells allows the production potential  
 CC of the microorganisms to be improved. New genes may be discovered,  
 CC possible functions of unknown open reading frames can be identified and  
 CC gene copy number variation and stability can be monitored. The expression  
 CC of genes can be used to study how FF cells adapt to changes in culture  
 CC conditions, environmental stress, spore morphogenesis, recombination,  
 CC metabolic or catabolic pathway engineering. Using ESTs provides several  
 CC advantages over genomic or random cDNA clones including elimination of  
 CC redundancy as one spot on an array equals one gene or open reading frame,  
 CC and organization of the microarrays based on function of the gene  
 CC products to facilitate analysis of the results. AAF07478 to AAF11247  
 CC represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents  
 CC ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from  
 CC Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from  
 CC Trichoderma reesei, which are all specifically claimed in the present  
 CC invention  
 XX  
 SQ Sequence 1146 BP; 262 A; 313 C; 274 G; 297 T; 0 U; 0 Other;

Query Match 9.1%; Score 74.2; DB 3; Length 1146;  
 Best Local Similarity 50.6%; Pred. No. 6.6e-11;  
 Matches 206; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 349 GATTACCCAGCAAGACCGCTGAGAGATGTGAGATTTGCTTGTGAT 408  
 DB 567 GAGTATAGCGCCAGGACGCCAAGACGATGTTGCAAGTCACTGCTGTTCATG 626  
 QY 409 GTTTCAGAGCTTTGCTTAAGCATTTGATCAAGAGATATCAAGGCTCTTCTGTTT 468  
 DB 627 ACTTCAGAGCGCTTGTCTTAAGCAATGATTCCTTCG---CAATGAGGTGACATCGCA 683  
 QY 469 TTGATTGTTTATGTCGTGCTGATTCGATTCGATTCCTCAAAACCAAGTTGCTACAC 528  
 DB 684 CTATTTGCGAGCATGAGTGTGCTATTTGCGAATCGGGGCTTTATCTGCCCTGCTTACAT 743

QY 529 ATGTCCAGGCTGTGTTATTCATTTGCTTAAGACTTTGGCTTGTGAATGGGCTAAGTAC 588  
 DB 744 GTTACCAAGGCTGAGTGCTTCAACTTCCGTTACCTGCGATGAGTGGGGCCGTTAC 803  
 QY 589 AACATCAAGTTAATTTCTTTAAACCCAGTTACATCTACGTTCTTTGAACCAAGATGTT 648  
 DB 804 AACATTGACATCAACACCATCTCGCCGCTACATTTGATGCTCATGTTGAGAGCTC 863  
 QY 649 ATCAATGTGAACGAAGAAATTGTAACAAGATGATCTGGTATCCCAACAAAGATG 708  
 DB 864 TTCTGTGATTTCTTGAAGCTCGCGAAGAAATGCCCCAAACATATATCTGGAGCTCTG 923  
 QY 709 TCCGACCAAGAAAGAAATACATTGCTGCTTTTGTATCTTCTTCTGA 755  
 DB 924 TCTACCCCTTAACGATACCGTGGCGCTGCGCTCTTCTTCAAGTGA 970  
 RESULT 8  
 ADS47380  
 ID ADS47380 standard; cDNA; 783 BP.  
 AC ADS47380;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Bacterial polynucleotide #2123.  
 XX  
 KM Recombinant DNA construct; transformed plant; improved plant property;  
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KM pathogen tolerance; pest tolerance; plant disease resistance;  
 KM cell cycle pathway modification; plant growth regulator;  
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 XX bacterial polynucleotide; gene; ss.  
 XX  
 OS Bacteria.  
 XX  
 PN US2003233675-A1.  
 XX  
 PD 18-DEC-2003.  
 XX  
 PF 20-FEB-2003; 2003US-00369493.  
 XX  
 PR 21-FEB-2002; 2002US-0360039P.  
 XX  
 PA (CAOY/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.  
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 XX  
 DR WPI; 2004-061375/06.  
 XX  
 PT New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 XX  
 PS Claim 1; SEQ ID NO 25810; 122pp; English.  
 XX  
 CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polynucleotide used in  
 CC the scope of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification but was obtained in electronic  
 CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 783 BP; 193 A; 186 C; 173 G; 231 T; 0 U; 0 Other;

Query Match 8.3%; Score 67.6; DB 13; Length 783;  
 Best Local Similarity 51.2%; Pred. No. 4.7e-09;  
 Matches 213; Conservative 0; Mismatches 194; Indels 9; Gaps 2;

QY 171 CAAATGCGTACGAGAAGTGAAGTTGAAGAAGTCCAAAGATGGTTCATATGCTG 230  
 DB 171 CGAAAGGCTGCAAGAAATGCCCCAGCTAATGGTCAAGACTTATACATGCAAGT 230  
 QY 231 TGATATTTCTGATTCGATACCGTCAAGGTGTTGCTCAAGTTGCTAAGATTGG 290  
 DB 231 CGATGTCATATTCCTTAAGAGGTAGAACATGCTTGTGATTCAAAGCTCTTGA 290  
 QY 291 TAAAGTCCATTCAGCTGTTAACAAGCTGTTAAGTGAAGTCTGTAAGTCCATGTAAGA 350  
 DB 291 TAGATATGATATGTTGTCGCCAACAAGGTATTCACATGCGACGCGCAT--GA 347  
 QY 351 TTTCCAGCAAGACGCTGAGAAGATGGAAGTGAAGTGAAGTGGTCTTGTATATG 410  
 DB 348 CATGACTTACGAAGATTGCTAATGAATAAAGTCAACTGCTGGTGTCTTCAACGT 407  
 QY 411 TTCTCAAGCTTGTCTAAGCATTCATCAAGAGGTATCAAGAGGTCTTCTGTTT 470  
 DB 408 CGGTCAAGATGCGGCCCATCTTCCAAAGCAAG-----CATGCGACGCTTGTAG 461  
 QY 471 GATTGCTTCAATGCTGATGCAATGTCATCAAGATCTCAAAACCAAGTTGTCTTAAACAT 530  
 DB 462 TACTGCTTCATGCTGATGTTGTGTAAAGTCTCAACACATGCTTACAAAC 521  
 QY 531 GTCCAGGCTGCTGTTATCATTTGGCTAAGACTTTGGCTTGTGATGGCTAAGT 586  
 DB 522 TTCCAAGCGCGGTGATCACTCATCAAGAGTTGGCAGTGAATGGCGCAAT 577

# RESULT 9

ADN62512  
 ID ADN62512 standard; DNA; 789 BP.

XX ADN62512;

DT 01-JUL-2004 (first entry)

XX A. thaliana gene SEQ ID NO:17.

XX ds; gene; seed size; organ size; plant; transgenic.

OS Arabidopsis thaliana.

XX Key Location/Qualifiers

FT CDS 1..789

XX WO2003096797-A2.

XX 27-NOV-2003.

XX 14-MAY-2003; 2003WO-US014989.

XX 15-MAY-2002; 2002US-038110P.

XX (MONS ) MONSANTO TECHNOLOGY LLC.

XX P1 Fernandes M, Xie Z, Dotson SB,

XX WPI; 2004-042481/04.

XX P-PDSB; ADN62513.

PT Increasing seed and organ size of a plant by transforming the plant with

PT a DNA construct comprising a promoter that functions in plants and

PT selecting a desired plant from a population of transformed plants

XX containing the DNA construct.

XX Claim 1; SEQ ID NO 17; 71pp; English.

CC The invention relates to a novel method for increasing seed and organ  
 CC size of a plant by: transforming the plant with a DNA construct  
 CC comprising a promoter that functions in plants, operably linked to a DNA  
 CC molecule that encodes a protein; and selecting a desired plant from a  
 CC population of transformed plants; containing the DNA construct; where the  
 CC desired plant exhibits increased seed and organ size compared to a plant  
 CC of a same plant species not transformed to contain the DNA construct. The  
 CC method of the invention is useful for increasing seed and organ size of a  
 CC plant. The present sequence is used in the exemplification of the  
 CC invention.

XX Sequence 789 BP; 220 A; 151 C; 183 G; 235 T; 0 U; 0 Other;

Query Match 8.1%; Score 66.4; DB 12; Length 789;  
 Best Local Similarity 56.5%; Pred. No. 1.1e-08;  
 Matches 165; Conservative 0; Mismatches 121; Indels 6; Gaps 2;

QY 519 TGTCTACAAATGTCACAGGCTGCTGTTATCCATTTGGCTAAGCTTTGGCTTGAATG 578  
 DB 471 TATCTATAGTTTACGAAAGAGCTTGAATCAATCAAGCAAAACCTTGGCATGTGAATG 530  
 QY 579 GGCTAAGTCAACATCAGATTAATCTTAAACCAAGTTACATCAAGTCTTGTGAC 638  
 DB 531 GGCAGAGATGATGATAGAGCCAACTGTTGCTCTTAATTTTATCTAAGTCTTATGGC 590  
 QY 639 CAAGATGTTATCAATGTTAAGCAAGATTTGACAAAGATGATCTGTTATCCACA 698  
 DB 591 TCAACTTTTTCAAAG---ACGCCATTACAGAAAGTTTGGTTAGAGAACTCACT 647  
 QY 699 ACAAGATGTCGCAACCAAGAAATACATGCTGCTGTTTGTATCTGTTTGAATC 758  
 DB 648 TGTGCGCGCTGAGAGCCAAATGAGG---TTTCATCACTTGTGGCTTTCTGTGTAC 704  
 QY 759 TGTGCTTATACATCACTGCTGCTGCACTTACTGTTGATGATGCTTACT 810  
 DB 705 TGCAGCTTCAATATTACTGTGACCACTTTGTGTGATGAGGCTTACT 756

# RESULT 10

ADS48180  
 ID ADS48180 standard; cDNA; 774 BP.

XX ADS48180;

DT 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #2923.

XX Recombinant DNA construct; transformed plant; improved plant property;

XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

XX pathogen tolerance; pest tolerance; plant disease resistance;

XX cell cycle pathway modification; plant growth regulator;

XX homologous recombination; seed oil yield; protein yield; carbohydrate;

XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

XX bacterial polynucleotide; gene; ss.

XX Bacteria.

PN US200323675-A1.  
XX 18-DEC-2003.  
XX 20-FEB-2003; 2003US-00369493.  
XX 21-FEB-2002; 2002US-0360039P.  
XX (CAOY/) CAO Y.  
XX (HINK/) HINKLE G J.  
XX (SLAT/) SLATER S C.  
XX (CHEN/) CHEN X.  
XX (GOLD/) GOLDMAN B S.  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
XX for expression of a polynucleotide encoding a polypeptide from a  
XX microbial source, useful for producing plants with improved properties.  
XX  
XX Claim 1; SEQ ID NO 26610; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
XX promoter functional in a plant cell, where the promoter is positioned to  
XX provide for expression of a polynucleotide encoding a polypeptide from a  
XX microbial source. The invention also relates to a transformed plant  
XX comprising the recombinant DNA construct and a method of producing a  
XX transformed plant having an improved property. The plant is a crop plant  
XX such as maize or soybean. The method of producing a transformed plant  
XX having an improved property comprises transforming a plant with the  
XX recombinant DNA construct and growing the transforming plant, where the  
XX polynucleotide or polypeptide is useful for improving plant properties.  
XX The recombinant DNA construct is useful for producing plants with  
XX improved plant properties, e.g. improved cold, heat or drought tolerance,  
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
XX increased resistance to plant disease, better growth rate by modification  
XX of the cell cycle pathway with plant growth regulators, increased rate of  
XX homologous recombination, modified seed oil or protein yield and/or  
XX content, improved yield by modification of carbohydrate, nitrogen or  
XX phosphorus use and/or uptake, by modification of photosynthesis or by  
XX providing improved plant growth and development under at least one stress  
XX condition, improved lignin production or improved galactomannan  
XX production. This sequence represents a bacterial polynucleotide used in  
XX the scope of the invention. Note: The sequence data for this patent did  
XX not form part of the printed specification but was obtained in electronic  
XX format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
XX Sequence 774 BP; 226 A; 160 C; 229 G; 159 T; 0 U; 0 Other;  
SQ  
Query Match 7.8%; Score 63.6; DB 13; Length 774;  
Best Local Similarity 61.4%; Pred. No. 6.9e-08;  
Matches 102; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
QY 472 ATTGGTTATGTCGTGGTCCATTGTCAACGATCTCTGAGCGGCGCCGAGTGGCCCGGTACCGA 531  
DB 436 ATCGGCTCATGTCCGGACACATCGTGAACAAACCTCAAGACAGACGTTACCAACGCT 495  
QY 532 TCCAAAGCTGTGTTATTCATTGGCTAAGACTTTGGCTTGGATGGGCTTAAGTACAC 591  
DB 496 TCGAAAGCGGTGTATCATCTCAACGATCTCTGCGCGCCGAGTGGCCCGGTACCGA 555  
QY 592 ATCAGAGTTAATCTTTAAACCCAGGTTACATCTACGGTCTTGA 637  
DB 556 ATCAGGGTAAACGATTAAGCCCGGATTCATCAAGAACCTCTCA 601  
RESULT 11  
AAFI1343  
ID AAFI1343 standard; cDNA; 636 BP.  
XX  
XX AAFI1343;  
AC

XX 13-MAR-2001 (first entry)  
XX Aspergillus niger EST SEQ ID NO:3866.  
XX  
XX Multiple gene expression; filamentous fungal cell; EST;  
XX expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
XX Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
XX culture condition; environmental stress; spore morphogenesis;  
XX metabolic pathway engineering; catabolic pathway engineering; ss.  
XX Aspergillus niger.  
XX WO200056762-A2.  
XX  
XX 28-SEP-2000.  
XX  
XX 22-MAR-2000; 2000WO-US007781.  
XX  
XX 22-MAR-1999; 99US-00273623.  
XX  
XX (NOVO ) NOVO NORDISK BIOTECH INC.  
XX (NOVO ) NOVO NORDISK AS.  
XX  
XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
XX WPI; 2000-594572/56.  
XX  
XX Monitoring differential expression of genes in filamentous fungal cells  
XX PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
XX substrate of expressed sequence tags.  
XX  
XX Claim 87; Page 1728; 3161pp; English.  
XX  
XX The present invention describes a method for monitoring differential  
XX expression of genes in a first filamentous fungal (FF) cell relative to  
XX expression of the same genes in one or more second filamentous fungal  
XX cells. The method uses fluorescence-labeled nucleic acids isolated from  
XX the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
XX are used in the methods for monitoring differential expression of genes  
XX in a first filamentous fungal (FF) cell relative to expression of the  
XX same genes in one or more second filamentous fungal cells. Monitoring the  
XX global expression of genes from FF cells allows the production potential  
XX of the microorganisms to be improved. New genes may be discovered,  
XX possible functions of unknown open reading frames can be identified and  
XX gene copy number variation and stability can be monitored. The expression  
XX of genes can be used to study how FF cells adapt to changes in culture  
XX conditions, environmental stress, spore morphogenesis, recombination,  
XX metabolic or catabolic pathway engineering. Using ESTs provides several  
XX advantages over genomic or random cDNA clones including elimination of  
XX redundancy as one spot on an array equals one gene or open reading frame,  
XX and organisation of the microarrays based on function of the gene  
XX products to facilitate analysis of the results. AAF07478 to AAF11247  
XX represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents  
XX ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from  
XX Trichoderma reesei, which are all specifically claimed in the present  
XX invention  
XX  
XX Sequence 636 BP; 169 A; 155 C; 161 G; 149 T; 0 U; 2 Other;  
SQ  
Query Match 7.4%; Score 60.6; DB 3; Length 636;  
Best Local Similarity 51.3%; Pred. No. 4.7e-07;  
Matches 194; Conservative 0; Mismatches 175; Indels 9; Gaps 2;  
QY 211 AAGATGGTTATATGCTGTGATATTTCTGATCTGTACCGTTACAGAGTGTGCT 270  
DB 200 AAGTACGCTGCCTTAAGTATGATGTCCGAATCAGAGATATTTAGAGAGGTATCCAG 259  
QY 271 CAAGTTGCTAAGATTTTGTGATGTGCATTCGATCTGTTACACAGCTGTATACGT 330  
DB 260 CAGATTTCTCAGACTTTGGCAAGCTTGATATTTATCTGTGGAAC--TCGGGGGTCACT 316







PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-01300449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134841P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135133P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 17-JUN-1999; 99US-0139453P.  
 PR 18-JUN-1999; 99US-0139452P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140981P.  
 PR 30-JUN-1999; 99US-014187P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142820P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145813P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.

```
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-015938P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 6.9%; Score 56.4; DB 3; Length 798;

Best Local Similarity 54.5%; Pred. No. 8.8e-06; Matches 158; Conservative 0; Mismatches 126; Indels 6; Gaps 2;

```
QY 521 TCTACAACATGTCACAGGCTGTGTTATTCATTGCTAGACCTTGGCTTGATGGG 580  
DB 482 TTTATAGTCTAATAAGAGGCTTGAACAGTTAAGTAATTTAGCATGTGATGGG 541  
QY 581 CTAAGTACAACATCAGAGTTAATCTTTPAACCCAGGTTACATCTACGGTCTTGACCA 640  
DB 542 CAAAAGACGGCATTAAGAGCCAAATGCTGTGCACCTAATGTGTCAAGACTCCTTGTCTTC 601  
QY 641 AGAATGTTATCAATGTTACGAAGAATTGTACACAGATGATCTGTGGTATCCACAAC 700  
DB 602 AATCTTATTCGAGG---ACGTGGGTTTCAAGAGGCAATTGTTCAATAGACTCCACTTG 658  
QY 701 AAAGATGTCGAAACCAAGAAATACATTGTGCTGTTTGTACTGTGCTTCTGAATCTG 760  
DB 659 GTGGGCTGAGAGCGGATGAAG---TTGCATCACTAGTGGCTTCTTGTGTACACAG 715  
QY 761 CTGCTTCAATCACTACTGTGCTGCCAGCTTACTGTTGATGTGTTCACT 810  
DB 716 CAGCTTCTTATTAATCTGTGTCAACCATTTGTATTGATGTGTGTTCACT 765
```

Search completed: March 2, 2005, 14:53:06  
Job time : 564 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 2, 2005, 05:24:38 ; Search time 3855 Seconds  
(without alignments)  
10256.668 Million cell updates/sec

Title: US-10-720-018-1

Perfect score: 816

Sequence: 1 atgactgactacattcccaac.....atgctggtcttccattcttgg 816

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_hlg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_scs:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	816	100.0	816	AJ583159	AJ583159 Ambrosioz
2	332.8	40.8	1166	CTARDH	U00675 Candida tro
3	305.6	37.5	1614	YSARDH	L16227 Candida alb
4	275.2	33.7	110000	CR382133_10	Continuation (11 o
5	268.2	32.9	1157	PSARDH	246866 P. stipitib
6	238.2	29.2	696	AR031556	AR031556 Sequence
7	238.2	29.2	696	AR345028	AR345028 Sequence
8	218.6	26.8	846	AR544898	AR544898 Sequence
9	155.8	19.1	110000	CR382121_09	Continuation (10 o
10	130.8	16.0	110000	CR382132_03	Continuation (4 of
11	92.6	11.3	110000	CR382133_11	Continuation (12 o
12	91.4	11.2	110000	CR382137_06	Continuation (7 of
13	87.2	10.7	4293	AF002134	AF002134 Candida a
14	86.8	10.6	126105	AC105434	AC105434 Megaport
15	86.6	10.6	927	AR547359	AR547359 Sequence
16	81.8	10.0	1485	AY387591	AY387591 Emerice11
17	81.8	10.0	15779	AY434691	AY434691 Emerice11
18	78.6	9.6	110000	CR382135_21	Continuation (22 o
19	78.6	9.6	110000	CR382131_06	Continuation (7 of

c	20	72	8.8	110000	1	AE016827_14	Continuation (15 o
c	21	70.8	8.7	110000	8	CR382133_10	Continuation (11 o
c	22	70	8.6	1063	8	UF081790	U81790 Uromyces fa
c	23	68.6	8.4	409	11	KLAJ9366	AJ29366 Kluyverom
c	24	68.2	8.4	340900	1	SMES91791	AL591791 Sinorhizo
c	25	67.6	8.3	32814	8	SPAC22A12	Z99295 S.pombe chr
c	26	64	7.8	6430	8	AFU320167	AJ320167 Aspergill
c	27	64	7.8	110000	8	CR382138_00	CR382138 Debaryomy
c	28	63.6	7.8	10126	1	AE001711	AE001711 Thermotog
c	29	62.4	7.6	309135	1	AP005946	AP005946 Bradyrhiz
c	30	61.8	7.6	873	6	AR547523	AR547523 Sequence
c	31	61.4	7.5	110000	8	CR382132_02	Continuation (3 of
c	32	61	7.5	786	8	STU245634	AJ245634 Solanum t
c	33	59.8	7.3	822	6	AX772916	AX772916 Sequence
c	34	59.8	7.3	1129	8	DURKROPI	L20473 Datura stra
c	35	59.4	7.3	15350	8	SPAC8E11	AL021817 S.pombe c
c	36	59	7.2	774	6	E30144	E30144 Arabiditol
c	37	59	7.2	2255	6	E30145	E30145 Arabiditol
c	38	57.8	7.1	300787	1	AE016744	AE016744 Staphyloc
c	39	57.2	7.0	37869	8	CR382137_20	Continuation (20 o
c	40	57.2	7.0	110000	8	CR382137_19	Continuation (20 o
c	41	56.6	6.9	976	8	STU400815	AJ400815 Solanum t
c	42	56.4	6.9	789	8	BT010425	BT010425 Arabidops
c	43	56.4	6.9	825	6	AX772918	AX772918 Sequence
c	44	56.4	6.9	909	8	AK175221	AK175221 Arabidops
c	45	56.4	6.9	1099	8	D88156	D88156 Hyoscyamus

## ALIGNMENTS

RESULT 1	AJ583159	816 bp	mRNA	linear	PLN 13-APR-2004
LOCUS	AJ583159				
DEFINITION	Ambrosiozyma monospora partial mRNA for NADH L-xylose reductase (aix1 gene).				
ACCESSION	AJ583159	GI:42820317			
VERSION	aix1 gene; NADH L-xylose reductase.				
KEYWORDS	Ambrosiozyma monospora				
SOURCE	Ambrosiozyma monospora				
ORGANISM	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycopsidaceae; Ambrosiozyma.				
REFERENCE	1 Verho, R., Putkonen, M., Londeborough, J., Penttila, M. and Richard, P. A Novel NMDH-linked L-xylose Reductase in the L-Arabinose Catabolic Pathway of Yeast				
AUTHORS	J. Biol. Chem. 279 (15), 14746-14751 (2004)				
JOURNAL	J. Biol. Chem. 279 (15), 14746-14751 (2004)				
PUBMED	14736891				
REFERENCE	2 (bases 1 to 816)				
AUTHORS	Richard, P.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-SEP-2003) Richard P., VTT Biotechnology, VTT, PO box 1500, Tietoele 2, Espoo, 02044 VTT, FINLAND				
FEATURES	Location/Qualifiers				
source	1..816				
gene	/organism="Ambrosiozyma monospora"				
CDS	/mol_type="mRNA"				
	/strain="NRRL Y-1484"				
	/db_xref="taxon:43982"				
	1..816				
	/gene="aix1"				
	1..>816				
	/gene="aix1"				
	/codon_start=1				
	/product="NADH L-xylose reductase"				
	/protein_id="CAE47547.1"				
	/db_xref="GI:42820318"				
	/translation="MTDTIPTRPFGHLLTYTGACGGLAEALIKLLAYGSDIALDI				
	DOETPAKQAEYHKAETELKLEPKRGSYACDISDVTAKVNAQVADGKPLH				
	LVNTAGYCEPCEDEYPAKNAEKMKVNLGSLVYSOAFKFLIEGKIGKASVNVIS				
	MSGALVNDPQNVVNVMSKAGVYHIAKTLACWAKYKRVNSLNGGYIIGPLTKVIV				
	GNEELYNMISGIPQGRMSPEKYEYIGAVLVLLSEASAYTTGASLLVDCGPTSW"				

## ORIGIN

Query Match 100.0%; Score 816; DB 8; Length 816;

Best Local Similarity 100.0%; Pred. No. 6.3e-196; Mismatches 0; Indels 0; Gaps 0;

Matches 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACTGACTACATTCACATCTTTAGATTCATGCGCCACTTACCATTCACAGTGC 60  
 DB 1 ATGACTGACTACATTCACATCTTTAGATTCATGCGCCACTTACCATTCACAGTGC 60  
 QY 61 TGTGTGCTTTAGCTGACAGCTTTATCAAGGCTTTGTGGCTTACGCTTCTGACATGCT 120  
 DB 61 TGTGTGCTTTAGCTGACAGCTTTATCAAGGCTTTGTGGCTTACGCTTCTGACATGCT 120  
 QY 121 TTGCTTGATATGACCAAGAAAGACTGCTGCCAACAAGCCGAATACCAAAATACGCT 180  
 DB 121 TTGCTTGATATGACCAAGAAAGACTGCTGCCAACAAGCCGAATACCAAAATACGCT 180  
 QY 181 ACTGAAAGATTTGAAGTTGAAAGATTCGAAAGATGGGTTCAATATGCTGTGATATTTCT 240  
 DB 181 ACTGAAAGATTTGAAGTTGAAAGATTCGAAAGATGGGTTCAATATGCTGTGATATTTCT 240  
 QY 241 GATTCTGATACCGCTTCAACAAGCTTTGCTCAAGTCTTCAAGATTTTGGTGAATGCCA 300  
 DB 241 GATTCTGATACCGCTTCAACAAGCTTTGCTCAAGTCTTCAAGATTTTGGTGAATGCCA 300  
 QY 301 TTGCACTTGCTTAAACAAGCTGCTTACTGTGAAAACCTTCCATGTGAAGATTAACCGAGCC 360  
 DB 301 TTGCACTTGCTTAAACAAGCTGCTTACTGTGAAAACCTTCCATGTGAAGATTAACCGAGCC 360  
 QY 361 AAGAAGCGCTGAGAAAGATGCTGAAAGTTAACTTTGGGTTCTTTGATGTTTCTCAAGCC 420  
 DB 361 AAGAAGCGCTGAGAAAGATGCTGAAAGTTAACTTTGGGTTCTTTGATGTTTCTCAAGCC 420  
 QY 421 TTTCGTAAGCCATTTGATCAAGAAGTATCAAGGCTGCTTCTGTTGTTGATGTTGCTTCT 480  
 DB 421 TTTCGTAAGCCATTTGATCAAGAAGTATCAAGGCTGCTTCTGTTGTTGATGTTGCTTCT 480  
 QY 481 ATGCTGCTGATGCTGATCAAGATCTCTCAAAACCAAGTGTCTACAACAATGCCAAGCT 540  
 DB 481 ATGCTGCTGATGCTGATCAAGATCTCTCAAAACCAAGTGTCTACAACAATGCCAAGCT 540  
 QY 541 GGTGTTATTCATTTGGCTTAAGACTTTGGCTTGTGAATGGGCTAAGTACAACATCAAGATT 600  
 DB 541 GGTGTTATTCATTTGGCTTGAAGCTTTGGCTTGTGAATGGGCTAAGTACAACATCAAGATT 600  
 QY 601 AATTCTTTAAACCCAGGTTACATCAAGCTCTTTGACCAACAAGATTTTCAATGCTTAC 660  
 DB 601 AATTCTTTAAACCCAGGTTACATCAAGCTCTTTGACCAACAAGATTTTCAATGCTTAC 660  
 QY 661 GAAGAATTGTACAACAGATGATCTCTGATCCCAACAAGAATGCCAACAAG 720  
 DB 661 GAAGAATTGTACAACAGATGATCTCTGATCCCAACAAGAATGCCAACAAG 720  
 QY 721 GAATACATTTGATGCTGTTTGTACTGCTTTGATGCTGCTTCAATACCTACTGCT 780  
 DB 721 GAATACATTTGATGCTGTTTGTACTGCTTTGATGCTGCTTCAATACCTACTGCT 780  
 QY 781 GCCAGCTTACTGCTGATGCTGCTTCACTTTCTGG 816  
 DB 781 GCCAGCTTACTGCTGATGCTGCTTCACTTTCTGG 816

## RESULT 2

CTARDH 1166 bp DNA linear PLN 16-MAY-1995  
 LOCUS Candida tropicalis D-arabinitol dehydrogenase (ARD) gene, complete cds.  
 ACCESSION U00675  
 VERSION U00675.1 GI:392785  
 KEYWORDS Candida tropicalis  
 SOURCE Candida tropicalis  
 ORGANISM

## REFERENCE

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 1 (bases 1 to 1166)  
 Murray, J.S., Wong, M.L., Miyada, C.G., Switchenko, A.C., Goodman, T.C. and Wong, B.

## TITLE

Isolation, characterization and expression of the gene that encodes D-arabinitol dehydrogenase in Candida tropicalis

## JOURNAL

Gene 155 (1), 123-128 (1995)

## MEDLINE

95212917

## PUBMED

7698655

## REFERENCE

2 (bases 1 to 1166)

## AUTHORS

Wong, B.

## TITLE

Direct Submission

## JOURNAL

Submitted (10-AUG-1993) Brian Wong, Internal Medicine, University of Cincinnati College of Medicine, 231 Bethesda Avenue, Cincinnati, OH 45267-0560, USA

## FEATURES

## source

Location/Qualifiers

## organism

1..1166  
/organism="Candida tropicalis"

## mol\_type

/mol\_type="genomic DNA"

## strain

/strain="ATCC 750"

## db\_xref

/db\_xref="taxon:5482"

## clone

/clone="pCT1A"

## clone\_lib

/clone\_lib="genomic DNA library in pAB107 from V.F. Kalb and J.C. Lopez"

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

## CDS

241..1089

## gene

241..1089

Db 638 ATGCCGAAATATCATGAAAGTCATAGCTTATAGATCCCTTTACGTCACAAAGCTTTGG 697  
 Qy 425 CTAAAGCATTTGATCAAAAGAGTATCAAGGGCTCTCTGTTGTTTGTATGTTCTATGT 484  
 Db 698 CCAGACCTTTGATTCAAAA---TAACATGACCTGGCTCTATTAATCTTATGATGGTTCTATGT 754  
 Qy 485 CTGGGCGCATTTGCAACGATCCCTCAAAACCAAGTTCTTACCAAGTCCCAAGGCTGGG 544  
 Db 755 CTGGTACCATTTGCAACGATCCCAACCAAGTTCTTACCAAGTCCCAAGGCTGGG 814  
 Qy 545 TTATCATTTGGCTTAAGACTTTGGCTTGAATGGCTAAGTACACATCAAGTAAAT 604  
 Db 815 TCATTCATCTGGCCAGATCTTTAGCATGTGAATGGCCCAATTCACATCAAGTAAAT 874  
 Qy 605 CTTTAAACCCAGATTACATCTAGCGCTCTTGAACCAAGATTTATCAATGGTTAAAG 664  
 Db 875 CTTGTCGCCAGGTTATATTTGACTCCATTCAGCAAGAAATGTTATAGTGTCACTG 934  
 Qy 665 AATTGACCAAGATGATCTCTGTATCCCAACAAAGATGTCGCAACCAAGGAT 724  
 Db 935 AATATGAAGCAGAAATGGAAATCCAGATTCATGAAGAAAGGCGGCAAAAGAT 994  
 Qy 725 ACATTTGCTGCTTTTGTACTTCTTCTGAAATCTGCTTCAATCACTACATGTCGCA 784  
 Db 995 TTGCTGCTTCTATTTTGTACTTCTGCTTCTGATTTCTCTTCTTCACTACATGTCGCA 1054  
 Qy 785 GCTTACTGTTGATGCTGTTCTTCACTCTTGG 816  
 Db 1055 ATTGTTAGTGAATGCTGTTGTTAGCAATGTTGG 1086

RESULT 3

YSAARH  
 LOCUS YSAARDH 1614 bp DNA linear PLN 21-JUN-1995  
 DEFINITION Candida albicans D-arabinitol dehydrogenase (ArdH) gene, complete cds.

ACCESSION L16227  
 VERSION L16227.1 GI:295567  
 KEYWORDS ArdH gene; D-arabinitol dehydrogenase.  
 SOURCE Candida albicans  
 ORGANISM Candida albicans  
 BUKARYOTA; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Mitosporic Saccharomycetales; Candida.

REFERENCE  
 AUTHORS Wong, B., Murray, J.S., Castellanos, M. and Croen, K.D.  
 TITLE D-arabinitol metabolism in Candida albicans: studies of the biosynthetic pathway and the gene that encodes NAD-dependent D-arabinitol dehydrogenase  
 JOURNAL J. Bacteriol. 175 (19), 6314-6320 (1993)  
 MEDLINE 94012494  
 PUBMED 8407803

REFERENCE  
 AUTHORS Wong, B., Leeson, S., Grindle, S., Magee, B., Brooke, E. and Magee, P.T.  
 TITLE D-arabinitol metabolism in Candida albicans: construction and analysis of mutants lacking D-arabinitol dehydrogenase  
 JOURNAL J. Bacteriol. 177 (11), 2971-2976 (1995)  
 MEDLINE 95286472  
 PUBMED 7768790  
 REFERENCE  
 AUTHORS Wong, B.  
 TITLE Direct Submision  
 JOURNAL Submitted (18-MAY-1993) Brian Wong, Department of Internal Medicine, University of Cincinnati College of Medicine, OH 45267-0560, USA  
 COMMENT  
 FEATURES  
 source Original source text: Candida albicans (strain WO1) DNA. Location/Qualifiers  
 1..1614  
 /organism="Candida albicans"  
 /mol\_type="genomic DNA"  
 /strain="WO1"  
 /db\_xref="taxon:5476"  
 /map="6; segment O"

ORIGIN

Query Match 37.5%; Score 305.6; DB 8; Length 1614;  
 Best Local Similarity 62.9%; Pred. No. 1.3e-66;  
 Matches 511; Conservative 0; Mismatches 289; Indels 12; Gaps 2;

gene /cissue\_lib="PEMBL23 from P.T. Magee"  
 435..1280  
 /gene="ArdH"  
 CDS 435..1280  
 /gene="ArdH"  
 /function="catalyzes D-arabinitol + NAD = D-ribose + NADH"  
 /codon\_start=1  
 /evidence=experimental  
 /transl\_table=12  
 /product="D-arabinitol dehydrogenase"  
 /protein\_id="AAC37430.1"  
 /db\_xref="GI:295568"  
 /translation="MDSAVMYVDNTPVSPFRDGLVILTSGSGIAVVRALAKGADVALDMLERTQCARPVLQGEOMKRPESPIGVYASNCIGDAEADLTPEKAI NEHKKISSVLNTVTVYAEFPAEYPAKEMENLKVGLSFVYSAFARPLIONM TGSILIGMSGTIVNDPOPCMVNMSVAGYIHLRLSLACEWAKNYINVTLSPEYIL TPLTNTVSGHTEMTWESKIPMRMEPKEFVGSIIYLASESASSYTHNLVVDG GYECW"

Db 14 TTCCAACTTTAGATTCAGTGGCACTTAACATTTGTCACAGTGGCTGTTGTTAG 73  
 Db 469 TCCCAAGCTTCGGTTGATGGAACAACTAGTATATTAACCGGTGCTCTGTTGTTGG 528  
 Qy 74 CTGAAGCTTTAATCAAGGGTTGTTGGCTTACAGTTCGACATTTGCTTGAATATG 133  
 Db 529 CTGCGGTGATCAAGAGCTTTATTAAGCAAGTGGCCGATTTGCTATTAATGATATA 588  
 Qy 134 ACCAAGAAAGACCTGCTGCAACCAAGCCGAATACCAAAATAGCTACTGAAGATGA 193  
 Db 589 ACTGGAAGAACACAAACAACTGCTAGAGAGCTTACAAATGGGGCAAGACAAATGA 648  
 Qy 194 AGTGAAGAAAGTTCCAA-----AGATGGGTTCATATCCCTGTATTTTCGANT 244  
 Db 649 AAGGTAAATACAAATACCAATGCGTACAGTGAAGTGTGTCATGTAATATTTGGAGT 708  
 Qy 245 CTGATACGCTTCAAGAGTGTGCTCAAGTTGCTTAAGATTTTGGTAAGTTCATTC 304  
 Db 709 CTGAAGCTGTCACTTTCATTAAGCAATCAAGAACACCAACGCAAAATCTCAAGT 768  
 Qy 305 ACTTGTTAACAACAGCTGTTTCTGTGAACCTTCCATGTGAAGATTAACCAAGCCACA 364  
 Db 769 TCTTGTTCAACACTGCGGTACGCTGAACCTTCCAGCTGAAGATGACCAAGCAAGA 828  
 Qy 365 AGCTGGAAGATGATGTAAGTTAATCTTGGGTCTTTTGTATGTTTCTCAAGCCTTG 424  
 Db 829 ACGTGAAGAACTGTGAAGATTAACGGTTGGGTGCTATTCAGTTTCCCAAGCTTTTG 888  
 Qy 425 CTAAAGCATTTGATCAAGAGGATATCAAGGGTCTTCTGTTTGTATGTTCTATGT 484  
 Db 889 CTAAAGCATTTATCCAAA---CAACATGACCGGATCATCTTTTGTATGCGGTCAAT 945  
 Qy 485 CTGGTGCATTTGTCAAGATCTCTCAAAAACCAAGTTGTCTAACAATGTCACAGGCTGG 544  
 Db 946 CCGGTACATGTCACAGCAACCAACCAACCAATGATGTACAAATGTCACAAAGCGGTG 1005  
 Qy 545 TTATCATTTGGCTTAAGACTTTGGCTTGAATGGCTTAAGTACCAATCAAGTAAAT 604  
 Db 1006 TCATTCATTTTACCGCATCATTTGGCTGATGATGAGGCTAAATCAATTCACAGTCAAC 1065  
 Qy 605 CTTTAAACCCAGATTACATCTAGCGTCTTGAACCAAGATTTTCAATGTTAAGCAAG 664  
 Db 1066 CATTTGCGCAGGATCATCTTAACCCATTAACAAGAACTTTATGATGACACACCG 1125  
 Qy 665 AATTGTACAAAGATGATCTGTGATTCACCAACAAAGATGTCGAACCAAGGAAT 724  
 Db 1126 AATATGAAGCAGAAATGGGAATCAAGATCCCAATGAAGAATGACAGAACCAAGAGAT 1185

QY 725 ACATTGTCGCTGTTTGTACTGCTTCTGAAATCTGCTGCTTCAATACACTAGTGCCCA 784  
 DB 1186 TTGTTGGTCTTATCTTATCTTACTTGAGGCTGAGAACTGCTTCAATACACTAGTGCCCA 1245  
 QY 785 GCTTACTGCTGATGCTGCTTCTTCACTTCTTGG 816  
 DB 1246 ACTTGCTGCTTGAACGGGGGCTTACGAGTCTGG 1277

RESULT 4  
 CR382133\_10/c  
 WPCOMMENT

Sequence split into 13 fragments LOCUS CR382133 Accession CR382133

Fragment Name	Begin	End
CR382133_00	1	110000
CR382133_01	100001	210000
CR382133_02	200001	310000
CR382133_03	300001	410000
CR382133_04	400001	510000
CR382133_05	500001	610000
CR382133_06	600001	710000
CR382133_07	700001	810000
CR382133_08	800001	910000
CR382133_09	900001	1010000
CR382133_10	1000001	1110000
CR382133_11	1100001	1210000
CR382133_12	1200001	1249565

Continuation (11 of 13) of CR382133 from base 1000001 (CR382133 Debaromyces hansenii ch

Query Match 33.7%; Score 275.2; DB 8; Length 110000;  
 Best Local Similarity 60.1%; Pred. No. 66-59;  
 Matches 492; Conservative 0; Mismatches 308; Indels 12; Gaps 2;

QY 14 TTCCAACTTTGATTCGATGCGCACTTAACTATGTCACAGTGCCTGCTGCTGTTAG 73  
 DB 28160 TTCCAACTTTGATTCGATGCGCACTTAACTATGTCACAGTGCCTGCTGCTGTTAG 28101  
 QY 74 CTGAAGCTTTAATCAAGGCTTTGTTGGCTTACGCTTCACTGATGCTTGTGTTGATTCG 133  
 DB 28100 CTGCGTGTGCTCTGAGAGCATTTATTAAGTCACAGGAGCAAAATGCTTGTGTTGATTCG 28041  
 QY 134 ACCAAGAAAGACTGCTGCCAAACACCCGAATACCAAAATACGCTACTGAAAGATTGA 193  
 DB 28040 ATTTAGAAAGAAACAAAGCTGCGAGCAAGAAATGGAATGCGGAGCAAAATCATTTGA 27981  
 QY 194 ACTTGAAGA-----AGTTCCAAAGATGGTTCAATGCTGATGCTGATTTCTGANT 244  
 DB 27980 AGGTCGACATGAAAGCCAGTGGAGCTGTTCCGATGCTGTTAATTTGGGATTT 27921  
 QY 245 CTGATACCGTTCAACAAGGCTGTTGCTCAAGTTGCTAAGATTTGGTAAGTGCATTGC 304  
 DB 27920 TTGAACAGGTGAGAGTGCTTTTAAACAATGAAACACAGATATGTCGTCGATTT 27861  
 QY 305 ACTTGGTTAACAAGCTGCTGTTACTGTTGAAACTTCCCATGTAAGATTAACCAAGCAAGA 364  
 DB 27860 TGTATATCAATCTGACAGGTGCTGTAAGAAATTTCCCGCCCAAGATATCCCATCTGCCA 27801  
 QY 365 AGCGTGAAGAGATGAGATGTTAACTTGTGGGTTCTTGTATGTTTCTCAAGCCTTGG 424  
 DB 27800 ATGCTGAAGATTTTGAAGGTCATGATGATTAAGTGATTTCTACGATATACAAGCATTTG 27741  
 QY 425 CTAAAGCATTTGATCAAGAAAGATATCAAGGCTGCTTCTGTTGTTGTTGATTTGTTCTANT 484  
 DB 27740 CCCGTCATGATCA---GCACAAATTAAGAAAGATTCATCATTTTAAAGTGGGTGATTT 27684  
 QY 485 CTGTCGCACTTGTCAACGATCTCTCAAAACCAAGTTGTCTAACAACATGTCACAGCTGCTG 544  
 DB 27683 CTGGAACAATTTGTGAACGACCAACCAACATGATATGTAACAACATGTCACAGCGAGTG 27624  
 QY 545 TTATCATTTGGCTAGACTTTGGCTTGTGATGGCTGTAAGTCAACATCAAGTAAATTT 604  
 DB 27623 TTATTCATTTGAACCAAGATCATTTGGCTTGTGAATGGCCAAATTTCAACATCAAGATTAACA 27564

QY 605 CTTTAAACCAAGTTTACATCTACAGCTCTTTGACCAAGAAATGTTATCAATGTTACGAAG 664  
 DB 27563 CTTTAAAGTCTGAGATATATCTTAAACCCCTTTGACCAAGAAAGTCAATTTCTGTCATTCGG 27504  
 QY 665 AATTGTACACAGATGATCTCTGATATCCCAACAAAGAAATGTCGCAACCAAGAAAT 724  
 DB 27503 ATATGAAAGAAAGCATGAGAAATCAAAAGTCCCAATGAAAGATGCGCAACCAAGAAAT 27444  
 QY 725 ACATTGTCGCTGTTTGTACTGCTTCTGAAATCTGCTGCTTCAATACACTAGTGCCCA 784  
 DB 27443 TTGTTGGCTCAATCTTGAATCTTGGCCAGAGTCTTCTTGAATTAATCAAGGTCAACA 27384  
 QY 785 GCTTACTGCTGATGCTGCTTCTTCACTTCTTGG 816  
 DB 27383 ACTTACTGCTGATGCTGCTTCACTGCTGCTG 27352

RESULT 5  
 PSARDH  
 LOCUS  
 DEFINITION  
 P. stipitis ARDH gene encoding D-arabitol dehydrogenase.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

Buxaricola; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; Pichia.  
 1 (bases 1 to 1157)  
 Hallborn, J., Wallfrieson, M., Penttila, M., Keranen, S. and  
 Hahn-Hagerdal, B.  
 A short-chain dehydrogenase gene from Pichia stipitis having  
 D-arabinitol dehydrogenase activity  
 Yeast 11 (9), 839-847 (1995)

JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

FEATURES  
 source  
 location/Qualifiers

gene  
 CDS  
 /organism="Pichia stipitis"  
 /mol\_type="mRNA"  
 /strain="CBS 6054"  
 /db\_xref="taxon:4924"  
 1..1157  
 /gene="ARDH"  
 191..1027  
 /gene="ARDH"  
 /function="NAD dependent, oxidises D-arabitol to  
 D-ribulose"  
 /evidence="experimental"  
 /product="D-arabitol dehydrogenase"  
 /protein\_id="CAA6939.1"  
 /db\_xref="GI:763164"  
 /db\_xref="GOA:P50167"  
 /translation="MDYSYANVNFRLDGLAITGSGGLAAVTSRALTAQADVA  
 LIDNMLERTKSAKEVIGWGEETLKGESASIGOVSAASCPIVGSAPLIONLRGS  
 HGKIALILINAGYCENPRAETVATNAESIKNVNGCSFVSGSPARLIONLRGS  
 IILGNSGTIVNDPQPCQTMKAGVTHVRSIACEMAKINIRVNTLSGYITPL  
 TRNVIISGHTKEKMESEKIPKMAEPKEFVGSLIYLASEASSYTTGHNLVGGEYE  
 CW"

# ORIGIN

Query Match 32.9%; Score 268.4; DB 8; Length 1157;  
 Best Local Similarity 60.1%; Pred. No. 3.66-57;  
 Matches 487; Conservative 0; Mismatches 311; Indels 12; Gaps 2;

```

QY 5 CTGACTACATCCACATTTAGATTCGATGCGCATTAACCATTTGTCACAGTGCCTG 64
DB 207 CTACGCTGTTCCCACTTCAGATTTGAGCGAAGATTGGCTATTATACCGAGGTTCTG 266
QY 65 GTGGTTAGCTGAAGCTTTAATCAAGGTTTGTGGCTTAAGGTTTGAATTCCTTGC 124
DB 267 GTGGTTAGCGGAGCATTTGCGTGCCTTGTGGCCGAGGCGGTGATGTTGCTCTCA 326
QY 125 TTGATATCGACCAAGAAAAGCTGCTCCCAACAGCCGAATACCAAAATACGCTAC 184
DB 327 TTGACATGAACCTTGAAGAAAGCAAGTCCGCTCCCAAGAAAGTTTGGGCTGGAG 386
QY 185 AAGAAATTGAAGTGAAGAAAGTCCAAAGATGGGTTCAATG-----CCTGATG 235
DB 387 AAGCTTTGAAGGTTGAAGCAAGCTTCAAGCATCGGCCAAGTTTCCGCTGCTTCAACA 446
QY 236 TTTCTGATTTCTGATACCGTTCAAGAGTGTTCCTAAGTGTGAAGATTTTGTAA 295
DB 447 TTGGGATGCTGAGGAGTGAAGCGTACTTTCAGCTTCATCAACGAACACAGGAG 506
QY 296 TGGCATTTGACCTGTTTAACACAGCTGTTTACTGTGAAAATTCCCATGTGAAGATTAC 355
DB 507 TCGCTGACTGTGTTAATACCGCTGAGTACTGTGAAAATTCCCTGCGAAAAGTAC 566
QY 356 CAGCCAAAGACGCTGAAGATGAGTAAAGTTAAGTTGGGTTCTTGTATGTTCTC 415
DB 567 CGGCTACTAAGCTTGAAGATCATGAAGTGAAGGTTTGGGCTAATTTACGTTTCCG 626
QY 416 AAGCTTTGCTAAGCATTCATCAAGAAAGTATCAAGGTTGCTTGTGTTTGAATG 475
DB 627 AATCGTTGCTAAGCATTCATCAAGAAAGTATCAAGGTTGCTTGTGTTTGAATG 683
QY 476 GTTCTATGCTGAGTGCATTTGTCAAGATCTCAAAACCAAGTTGCTACAAATGTCCA 535
DB 684 GCTCATGCTTGAAGAAATGTCAAGCAAGCAAGCAAGCAAGTATGTACAAATGTCCA 743
QY 536 AAGCTGATGTTATCCATTTGCTGCTAAGCTTTGCTGCTGCTGCTGCTGCTGCTG 595
DB 744 AAGCTGATGTTATCCATTTGCTGCTAAGCTTTGCTGCTGCTGCTGCTGCTGCTG 803
QY 596 GAGTTAATCTTTAAACCCAGTTTACATTCAGGCTCTTTGACCAAGAAATGTTATCA 655
DB 804 GAGTCAACACCTTATACCAAGGCTATATTTTGAATCTTTTAAACCAAGAAATGTTAT 863
QY 656 GTACCAAGAAATGTGAACAGATGATCTGTGTAATCCCAACAAAGATGTCCGAAC 715
DB 864 GCCACACAGAGTGAAGAAAGCTTGGGAATCCAAAGATCCCATGAAGAGATGCGGAA 923
QY 716 CAAGGAATACATTTGCTGCTTTTGTACTGCTTTCTGAATGCTGCTGCTTCAATCA 775
DB 924 CCAAGGAATTCGTTGGGCTCATTTAATCTTGGCAAGCAAGACTGCTTCTTCACTA 983
QY 776 CTGGTCCAGCTTACTGTTGATGATGTTG 805
DB 984 CGGCGCAATTTGTTGTTGAGAGGAGAT 1013

```

```

RESULT 6
LOCUS AR031556 696 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 6 from patent US 5866382.
ACCESSION AR031556
VERSION AR031556.1 GI:5945845
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 696)
AUTHORS Hallborn,J., Penttila,M., Ojamo,H., Walfredsson,M., Airaksinen,U.,
Keranen,S. and Hahn-Hagerdal,B.
TITLE Xylose utilization by recombinant yeasts
JOURNAL Patent: US 5866382-A 6 02-FEB-1999;
FEATURES Location/Qualifiers

```

```

source 1..696
/organism="unknown"
/mol_type="unassigned DNA"

```

```

Query Match 29.2%; Score 238.2; DB 6; Length 696;
Best Local Similarity 64.4%; Pred. No. 1.7e-49;
Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

```

```

QY 227 CCTGTGATTTCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 286
DB 107 CCTGCAACATTTGGGATGCTGAGGAGTGAAGCTTCACTTCACTTCACTTCACTT 166
QY 287 TTGTAAGTTGCAATTTGCACTTTGATTAACAGCTGTTTACGTGAAATCTTCCAT 346
DB 167 ACGCAAGATCGCTGATCTTTGATTAACAGCTGTTTACGTGAAATCTTCCAT 226
QY 347 AAGATTTCCCAAGCAAGCAAGCTGAGAAAGTGAAGTTAATTTGTTGGTTCTTTGT 406
DB 227 AAGATTTCCCAAGCAAGCAAGCTGAGAAAGTGAAGTTAATTTGTTGGTTCTTTGT 286
QY 407 AATTTCTCAAGCTTTGCTAAGCAATGATCAAGAAAGTATCAAGGTTCTTCTTGT 466
DB 287 AAGTTTCCCAAGCAAGCAAGCTGAGAAAGTGAAGTTAATTTGTTGGTTCTTTGT 466
QY 467 TTTGATTTGCTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526
DB 344 TCTGATTTGCTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
QY 527 ACATGTCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 586
DB 404 ACATGTCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 463
QY 587 ACAATCTCAAGTTAATTTTAAACCAAGTTAATTTTAAACCAAGTTAATTTTAAAC 646
DB 464 ACAATCTCAAGTTAATTTTAAACCAAGTTAATTTTAAACCAAGTTAATTTTAAAC 523
QY 647 TTAATCTCAAGTTAATTTTAAACCAAGTTAATTTTAAACCAAGTTAATTTTAAAC 706
DB 524 TGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 583
QY 707 TGTCCGAACCAAGAAATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 766
DB 584 TGTCCGAACCAAGAAATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643
QY 767 CATACCTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 805
DB 644 CCAATCTAAGTTAATTTTAAACCAAGTTAATTTTAAACCAAGTTAATTTTAAAC 682

```

```

RESULT 7
LOCUS AR345028 696 bp mRNA linear PAT 17-AUG-2003
DEFINITION Sequence 6 from patent US 6582944.
ACCESSION AR345028
VERSION AR345028.1 GI:33741148
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 696)
AUTHORS Hallborn,J., Penttila,M., Ojamo,H., Walfredsson,M., Airaksinen,U.,
Keranen,S. and Hahn-Hagerdal,B.
TITLE Production of ethanol from xylose
JOURNAL Patent: US 6582944-A 6 24-JUN-2003;
FEATURES Location/Qualifiers
source 1..696
/organism="unknown"
/mol_type="mRNA"
ORIGIN
Query Match 29.2%; Score 238.2; DB 6; Length 696;
Best Local Similarity 64.4%; Pred. No. 1.7e-49;

```

	Matches	373; Conserved	0; Mismatches	203; Indels	3; Gaps	1
QY	227	CCTGGAATATTTCTGATCTTGATACCGTTCAACAAGGTGTTGCTCAAGTGTCTGAATG				286
Db	107	CCTGGAACATTTGGGGATGTGTGAGGAGTAGACGCTACCTTTCAGTCCATCAACGAACCC				166
QY	287	TTGGTAAGTTGGCATTGCACTGGTTAACAAGCTGGTGTACTGTGTAACCTCCCATGTG				346
Db	167	ACGGCAAGATCGCTCACTTTGTTGATTTACAACCGCTGGATTACTGTAAACCTTCCCTGCCG				226
QY	347	AAGATTACCCCAAGCAAGAACGCTGAGAAAGATGTAAGGTTAACTGTTGGTCTTCTTGT				406
Db	227	AAACGTACCCGGCTACTAACAAGCTGAAGATCATGAAGTGAACGTTTGGGCTCATTCCT				286
QY	407	ATGTTTCTCAACGCTTTGTGCTAAGCATTTGATTAAGAAAGATATCAAGGTCCTTTGTTG				466
Db	287	ACGTTTCCCAATCGTTGGCTAGACCATTTGATCCAA--CAACTTGAAGAGGCTCTATCA				343
QY	467	TTTTGATTGGTCTTATGTCGTGGTGCATTTGCAAGATCCATAAACAAAGTGTCTTCA				526
Db	344	TCTTATTTGGCTCAATGTCTGGAACAATTTGTAAAGAACCCACAAACCCCAATGTATGTACA				403
QY	527	AACATGCCAAGGCTGGTGTATTCATTTGGCTAAGACTTTGGCTTTGTAATGGGCTAAGT				586
Db	404	ACATGTCCAAAGGCTGAGATGATCCACTTGGTCAAGTCGTGGCTCGCAATGGGCAAGT				463
QY	587	ACAACATCAAGTTAATCTTTAAACCAAGTTACATCTACGGTCCCTTGAACCAAGATG				646
Db	464	ACAAATCAAGTCAACACCTTATCAACAAGGCTATATTTTGACTCCTTTAACCAAAAACG				523
QY	647	TTATCAATGGTAAGCAAGATTGTACAAACAATGATCTCTGGTATCCCAACAACAAAGA				706
Db	524	TGATTTCTTGGCCACACAGAGATGAAGAAAGCTGGAAATCCAAAGATCCCATGAAGAGA				583
QY	707	TGTCCGAACCAAGGAATACATTTGGTGTCTTTTGTACTTGTCTTCTGAATCTGCTGCTT				766
Db	584	TGGCGGAACCCAAAGGAATTCGTGGGGTCCATCTTATTACTTGGCAAGCGAAGTCTTCTT				643
QY	767	CATCACTACTGTGTCCAGCTTACTCTGTTGTATGTGTCCTT				805
Db	644	CCTACACTACGGGCCACAATTTGGTTGTGACGAGAGAT				682

ORIGIN	Query Match:	26.8%;	Score 218.6;	DB 6;	Length 846;
	Best Local Similarity	62.0%;	Pred. No. 1.5e-44;		
	Matches 385;	Conservative	0;	Mismatches 224;	Indels 12; Gaps 2;
0y	14 TTCCAACCTTTAGATTCGATGCGCCACTTAACCATGTGTCAGAGTCCCTGTGGTGTAG	73			
Db	218 TCCCAAGCTTCGGTTGGATGGAALAACTAGTCATTTAACCGGAGGCTCTGGTGGTTGG	277			
0y	74 CTGAAGCTTTAATCAAGGCTTTGTGGCTACGGTTCTGAACATTGCTTGGCTTGATATCG	133			

Db	278	CTGCCGCGGTATCAAAAGCTTTATTAGCCAAAGGCCGAGTGTGCATTAATGATCAATATGA	337
Qy	134	ACCAAGAAAAGACTGCTGCCAAACAAAGCCGAATACCACAATATACGCTATCTGAAGATTGA	1939
Db	338	ACTTGGAAAGAACAAACAACAGCTGCTAGAGACGTTTCAATATGGGGCCGAAGAGCAAAATGA	397
Qy	194	AGTTGAAAAGAAAGTTCCAA-----AGATGGGTTCAATAGCCGTGATATTTCTGATTT	2444
Db	398	AAGGTAAATATCGAATATCAACATCGGTCAAGTGAAGTCTTGTCATGTATATATTTGGCGATG	457
Qy	245	CTGATACCGTTCAACAAGGTGTTTGCTCAAGTGTCTAAGAAATTTTGTAAAGTTGCATTGC	304
Db	458	CTGAAGCTGTGCACTTGACATTTCAAAAGCATCAAGAACACCAAGGCAAAATCTCAAGTG	51.7
Qy	305	ACTTGGTTAACACAGCTGGTTACTGTGAAAACCTTCCCATGTGAAGATTACCCAGCCAAAGA	364
Db	518	TCTTGGTCAACACTGCGGGTTACGCTGAAAATTTCCAGCTGAAGAGTACCACCAAGCAAGA	5777
Qy	365	ACGCTGAAGAGATGTGAAGTTTAACTTGTTGGGTTCTTGTATGTTTCTCAAGCCTTTG	4244
Db	578	ACGCTGAAAACCTTATGAAAGTTTAAACGGGTATGGGTCATTTCTACGTTTCCCAAGCTTTTG	6377
Qy	425	CTAAGCCATTGATCAAAAGAGTATCAAGGGTCTCTTGTTGTTTGAATGGTTCTATGT	4844
Db	638	CTAGACCATTAATCCAAAA---CAACATGACCGAGTCATCATTTTGTATGGGTCATATGT	6944
Qy	485	CTGTGTCAATGTGCAACGATCTCTCAAAACCAAGTTGTCTAACAATGTCCAAAGCGTG	5444
Db	695	CCGGTACATCGTCAACGACCAACAACCAATGCAATATCAATGCAATGTCCAAAGCCGGTG	7544
Qy	545	TTATTCATTTGGCTTAAGACTTTTGGCTTGTGATATGGGCTAAGTACAAATCAGATTAATT	6044
Db	755	TCATTCATTTAGCCAAATCATTTGGCTGTGATATGGGCTAATATCAATATCGAGTCAACA	81.44
Qy	605	CTTTAAACCGAGTTACATCT	625
Db	815	CATGTGCGCCCGGCAATCT	835

```

RESULT 9
CR382121_09/c
WPCOMMENT
Sequence split into 11 fragments LOCUS CR382121 Accession CR382121
Fragment Name Begin End
CR382121_00 1 110000
CR382121_01 100001 210000
CR382121_02 200001 310000
CR382121_03 300001 410000
CR382121_04 400001 510000
CR382121_05 500001 610000
CR382121_06 600001 710000
CR382121_07 700001 810000
CR382121_08 800001 910000
CR382121_09 900001 1010000
CR382121_10 1000001 1062590
Continuation (10 of 11) of CR382121 from base 900001 (CR382121 Kluveromyces lactis str
Query Match 19.1%; Score 155.8; DB 8; Length 110000;
Best Local Similarity 53.4%; Pred. No. 1.1e-28;
Matches 473; Conservative 0; Mismatches 337; Indels 75; Gaps 4;
QY 1 ATGACTGACATCATTTCCAACTTTTAAAGTTTGAGTGGCCACTTAACCATTTGTCACAGGTGCC 60
DB 84309 ATTACTGATGTTCTGCCTTCTTTTAAAGTTGATGGAACGATCACTCAATCGTCACTGGATGCT 84250
QY 61 TGTGATGTTTAAAGCTTAAAGCTTTTAAATCAAGGATTTTGTGGCTCAAGCTTCTGACATTTGCT 120
DB 84249 GCGGATGGGTTGGCTCACACTTTGTGTCAACGCGATTTGTGTCAAGGAGATCAGAACTAGCA 84190
QY 121 TTGCTTGATATCGACCAAGAAAAGACTGCTGCCAAACAGCGGATATCCACAATAATAGCT 180
DB 84189 TTGATGATATCTTCTCTTGGCTAGTTTAGAGATGTTAAAACTAATATGGAACTCTTC--- 84133

```





CR382133\_03 300001 410000  
 CR382133\_04 400001 510000  
 CR382133\_05 500001 610000  
 CR382133\_06 600001 710000  
 CR382133\_07 700001 810000  
 CR382133\_08 800001 910000  
 CR382133\_09 900001 1010000  
 CR382133\_10 1000001 1110000  
 CR382133\_11 1100001 1210000  
 CR382133\_12 1200001 1249565

Continuation (12 of 13) of CR382133 from base 1100001 (CR382133 Debaryomyces hansenii ch

Query Match 11.3%; Score 92.6; DB 8; Length 110000;  
 Best Local Similarity 52.5%; Pred. No. 1.1e-12;  
 Matches 309; Conservative 0; Mismatches 259; Indels 21; Gaps 4;

QY 221 CATATGCCGTGATATTTCTGATTCGATACCGTTCAAGAGTGTGCTCAAGTGTCTA 280  
 DB 99071 CTTATTAATGTCTGTTGATTAATTTGAAGAGTAGAAAAGCTGAATCAAAATAGTA 99012  
 QY 281 AGGATTTTGGTAAGTGGCATTGCACTTGATTAACAAGCTGTTACTGTGAAGACTTCC 340  
 DB 99011 AGGATTTTGGAAACAATGATTAATTTGCTGTAATGCTGGGTTGGCTGGAATCTGCC 98952  
 QY 341 CATGTGAAGATTACCCAGCCAAAGCGCTGAGAGAATGTGAAGTTAACTTTGGGTT 400  
 DB 98951 CAGTAAGCAGTCTTCAATAGATGCTTACCAACAATGTGATGAAGTTAACTTAAGACGGG 98892  
 QY 401 CTTTGTATGTTTCTCAAGCTTTGCTAAGCCATTGATCAAAAGGATCAAGGTTCTT 460  
 DB 98891 TTTTACTATGC-----AGCCAAATGTATAGACCAATTTTCAAAAGATGGGAAAGAT 98838  
 QY 461 CTGTTGTTTGAATGGTCTATATGCTGTGCCATTGTCAACGATCTGAAAACAAGTTG 520  
 DB 98837 CATTTATATTTACTGTTTCCCAAGCTGGACATATAGTCAAGTCTTATGCTCAAGCTG 98778  
 QY 521 TCTAACACATGTCAGAGGCTGTGTTATTCATTGGCTTAAGACTTTGGCTTGAATGG 580  
 DB 98777 CCTATATATGCAAGTAAGAGAGCTCTATCCAAATGGGAAATCACTAGCAATTAAGTGG 98718  
 QY 581 CTAAATGACACATCAGAGTTAATCTTTAAACCCAGTTTCACTAAGGTCCTTTGACCA 640  
 DB 98717 TTAGAT---TTGCAAGAGTTAATACTATATCAACAGTTAAT-----TTTGACCG 98670  
 QY 641 AGAATGTTTCAATGGTGAACGAAGATTGTACAAAGATGATCTGTGATCCCAAC 700  
 DB 98669 AGATGCGCAAAATGTTCCCTTTGATGTGAATGCAAAATGCGGCTTTAATTTCCAAAG 98610  
 QY 701 AAAGAATGTCGAAACCAAGAAATACATTGTGCTGTGTTTGTACTGCTTTCTGAATCG 760  
 DB 98609 GGAGAGAGGGCTGCGCAAGAAATAGTCGAGACATTTTGTATTTGCAATCGAC---G 98553  
 QY 761 CTGCTTCAATCACTAATGTTGCTGAGTTACTGCTGTGATGTGATGTTTCTGAC 809  
 DB 98552 CATTCATCACTCACTGCTGCTGAGCACTTAATATGATGCGGATATATAC 98504

RESULT 12  
 CR382137\_06  
 WPCOMMENT  
 Sequence split into 21 fragments LOCUS CR382137 Accession CR382137

Fragment Name	Begin	End
CR382137_00	1	110000
CR382137_01	100001	210000
CR382137_02	200001	310000
CR382137_03	300001	410000
CR382137_04	400001	510000
CR382137_05	500001	610000
CR382137_06	600001	710000
CR382137_07	700001	810000
CR382137_08	800001	910000
CR382137_09	900001	1010000
CR382137_10	1000001	1110000

CR382137\_11 1100001 1210000  
 CR382137\_12 1200001 1310000  
 CR382137\_13 1300001 1410000  
 CR382137\_14 1400001 1510000  
 CR382137\_15 1500001 1610000  
 CR382137\_16 1600001 1710000  
 CR382137\_17 1700001 1810000  
 CR382137\_18 1800001 1910000  
 CR382137\_19 1900001 2010000  
 CR382137\_20 2000001 2037969

Continuation (7 of 21) of CR382137 from base 600001 (CR382137 Debaryomyces hansenii chro

Query Match 11.2%; Score 91.4; DB 8; Length 110000;  
 Best Local Similarity 52.5%; Pred. No. 2.2e-12;  
 Matches 331; Conservative 0; Mismatches 276; Indels 24; Gaps 5;

QY 182 CTGAAGATTGAAGTTGAAGAAGTCCAAAGATGGTTCAATGCTTGAAGTATTTCTG 241  
 DB 92817 CTGAATACTTAACCAAGACCTACGAGTAGAAATCCAAAGCATACAAAGTATATTTCCG 92876  
 QY 242 ATTCTGATACGTTTCAAGAGTGTGCTCAAGTTGCTAAGGATTTGTAAGTGGCAT 301  
 DB 92877 ACCCAAGAGATGGAAGAAAGTATGATCGACAAATTAAGAAAGATTTCCGCAATTGACG 92936  
 QY 302 TGCACCTGTTAACAACAGCTGTTTACTGTGAAGAACTTCCATG--TGAAGATTACCCAG 358  
 DB 92937 TTTTGTGCTATATGCTGTGTGCTGATGAGACCGAGGAGAAATTAAGGCGATGGGT 92996  
 QY 359 CCAAGACGCTGAGAAATGTTAAGTTAACTGTTGGTCTTTGTATGTTTCTCAG 418  
 DB 92997 ACGATGCGTGAAGAAAGATGTAGATTTGAGACTTGAGTGAAGTTACTACCTGCGCTAAGA 93056  
 QY 419 CTTTGCTAAGCATTAATCAAGAAGATCAAGGATCAAGGATCTGTTGTTTGAATGGTT 478  
 DB 93057 CTGTGCGAAGATTTTCAAGAGCAGGAA-----AGGTTCTTGTGTATACAGCAT 93110  
 QY 479 CTATGCTGTGTCATGTTGCAAGATCTCAAAACCAAGTTGTCTACAAACATGTCAGAG 538  
 DB 93111 CAATGCTGGGCAATGTTGATGTGATGTTCCAAATGCAAGTCTCATCAACCGGTGTAAG 93170  
 QY 539 CTGCTGTTATCATTTGGCTTAAGACTTTGGCTTGTGAATGGGCTTAAGTCAACATCAG 598  
 DB 93171 CGGGGTATTCACATGTTGGTATGCTTGAAGTCTATGATGAGTGGCTTCATTCCT--AGAG 93227  
 QY 599 TTAATTTTAAACCCAGGTTATCATCTAAGGATCTTGACCAAGAAATTTATCATATGTA 658  
 DB 93228 TTAACACATTTCTTCAGATTACAT-----TGTACAGAGATTTCCGATTTTGTTC 93278  
 QY 659 ACGAAGATTGTACAAAGATGATCTCTGTATCCCAACAAGAAATGTCGGAACCA 718  
 DB 93279 CTGCGAAGTCAAGGCTAATATGTGCGCAATTGATCCATTGGAAGAGAACACTGCTC 93338  
 QY 719 AGGAATACATTTGCTGTTTGTACTGCTTTCTGAAATCTGCTGCTTCAATCACTACTG 778  
 DB 93339 AAGAGTTGTTGCTGCTTACTTAATTTTGTGATCGA---TGCCTTCACTTACACAGCTG 93395  
 QY 779 GTGCCAGTTACTGTTGATGTGTTTCTGAC 809  
 DB 93396 GCTGCGACTTAATGTTGATGTGTTGTTCTC 93426

RESULT 13  
 AF002134  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

AF002134 4293 bp DNA linear PLN 26-JUN-1998  
 Candida albicans Sou2p (S002), Sou1p (S001) and Yma8p (YMA8) genes,  
 complete cds.  
 AF002134  
 AF002134.1 GI:2183241  
 Candida albicans  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE 1 (bases 1 to 4293)  
 AUTHORS Janbon, G., Sherman, F. and Rustchenko, E.  
 TITLE Monosomy of a specific chromosome determines L-ascorbose utilization:  
 a novel regulatory mechanism in *Candida albicans*  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (9), 5150-5155 (1998)  
 MEDLINE 98226783  
 PUBMED 9560244  
 REFERENCE 2 (bases 1 to 4293)  
 AUTHORS Janbon, G., Rustchenko, E. and Sherman, F.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-MAY-1997) Department of Biochemistry and Biophysics,  
 University of Rochester, 601 Elmwood avenue P.O. Box 712,  
 Rochester, NY 14642, USA  
 location/Qualifiers  
 1..4293  
 /organism="Candida albicans"  
 /mol\_type="genomic DNA"  
 /strain="Sor17"  
 /db\_xref="taxon:5476"  
 /chromosome="4"  
 356..1199  
 /gene="SOU2"  
 356..1198  
 /gene="SOU2"  
 /note="short-chain alcohol dehydrogenase homolog"  
 /codon\_start=1  
 /transl\_table=12  
 /product="Soulp"  
 /protein\_id="AAC24462.1"  
 /db\_xref="GI:2183242"  
 /translation="MSKETTSYTNNAKGLPTKATIPNIIIDAFSLKGVASVTSS  
 GGIQAVAGYAGAGADVAIIVNSHPADKAEYLITGVKSKAYKCNTPDPVSKV  
 VKQIEDPFGTIDIFVANAAGVATREGKPFVKGVDKNNKVDVDDNSVYICAHVGEIF  
 RKXKGSFLPTASASIVNPQLOAVAKKAGVHLKSLFVEWAPFARVNSVSPG  
 YATHLSEFADPDKSKWQLPLGREAKREHVGAYLYIASDPAASYTGADLAVDGG  
 YTVI"  
 2063..2909  
 /gene="SOU1"  
 2063..2908  
 /gene="SOU1"  
 /note="short-chain alcohol dehydrogenase homolog"  
 /codon\_start=1  
 /transl\_table=12  
 /product="Soulp"  
 /protein\_id="AAC24463.1"  
 /db\_xref="GI:2183243"  
 /translation="MSEELISFTNPALGPLTPKAPOLPSNVLDLPSLKGKVASVTSS  
 GGIQAVAGYAGAGADVAIIVNSHPADKAEYLITGVKSKAYKCNTPDPVSKV  
 INIEKDPTGIDIFVANAAGVATREGKPFVKGVDKNNKVDVDDNSVYICAHVGEIF  
 KKKKSGSLITSSMSGTYINIPQLOAPYNAAKAACTHLAKSLSYEASFGARVNSISP  
 GYILIDIAFPADPEMKKMWQLPLGREGLPELVGAYLYIASDPAASYTGADLAVDGG  
 GYTC"  
 complement(3073..3877)  
 /gene="VMA8"  
 complement(3074..3877)  
 /gene="VMA8"  
 /note="Vacuolar H<sup>+</sup>-ATPase subunit 8 homolog"  
 /codon\_start=1  
 /transl\_table=12  
 /product="Vma8p"  
 /protein\_id="AAC24464.1"  
 /db\_xref="GI:2183244"  
 /translation="MSGAGNRECVPTTMTGLGVMSKLGKAGQGSGLLRKKEBALTKR  
 FDIYTRIDARKKMGVQTAFLAEVQYATGDIISYQVESVQKARFTYAKEN  
 VSGVFLPTDSHINEDVDFKALARGQGVQKALISKAVETLVEASLQTAFLI  
 LDEVKITRNRVNAIEHVIIPRENTIAYINGELIDEMDEEFLRKVOEKQENAAA  
 AEGEALAKAKAGATDELAIQDVADLIIKDKKEVDIIQEGEDVIV"  
 ORIGIN  
 Query Match 10.7%; Score 87.2; DB 8; Length 4293;  
 Best Local Similarity 49.7%; Pred. No. 2.9e-11;  
 Matches 314; Conservative 0; Mismatches 303; Indels 15; Gaps 3;

QY 182 CTGAGATTTGAAGTTGAAAGAGTTCAAGATGGTTCATATGCGCTGATATTCTG 241  
 |||||  
 Db 2286 CTGAATTTTAACTGAAAAATATGGTCTAAAGCCAAAGCTTTAATATGTAAGT 2345  
 QY 242 ATTTGATACCGTTCAAGAGTGTTCGTAAGTTGCTAAGATTATTTGGTAAGTCCAT 301  
 |||||  
 Db 2346 ATCTAATGATGTTTCTAAGATGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 2405  
 QY 302 TGCACTTGGTTAACAAGCTGGTACTGTGAAAACCTCCAGTGAAGATTACCCAGCCA 361  
 |||||  
 Db 2406 TATTTGTTCTATATGTGAGTTGACATGACATGACAAATATGTTCAAGGCT 2465  
 QY 362 AGAAGCTGAGAGATGAGAGGTTAACTTGTGGCTCTTGTATGTTTCAAGCCT 421  
 |||||  
 Db 2466 ATGATCAATGAAAAAGATGCTGATTTGATTTAATGAGTTTATTTGT--GCTC 2522  
 QY 422 TTGCTAAGCCATGATCAAGAGATGATCAAGAGGCTCTGTGTTTGTATGTTCTA 481  
 |||||  
 Db 2523 ATACGCTGGGACAAATCTTAAAGAAATTAATCTGTTCAATTAATTAATTAATTAATCA 2582  
 QY 482 TGTCTGTGCTCATTTGCAACATCTTCAAAACCAAGTTGTCACAAATGTCAGGCTG 541  
 |||||  
 Db 2583 TGTCAAGGCAATTTGTTAATATCCCTCAATTAACAGCTCTTATATGCTGCTAAAGCTG 2642  
 QY 542 GGTATTCATTTGGTGAAGCTTGGCTGTGAATGGCTAAGTACATCAGAGTTA 601  
 |||||  
 Db 2643 CATGTACTATTTAGCCAAATCATTTGATGATGATGATGATGATGATGATGATGATGAT 2702  
 QY 602 ATTCCTTAAACCCAGGTTACATTCAGGCTCTTGAACAAGATTTATCAATGATGATG 661  
 |||||  
 Db 2703 ATTCATATTTCTCAAGGCTATATA-----TTGATGATATTTGCTGATGATG 2753  
 QY 662 AGAATTTGACAACAGATGATCTGCTGATCCACACAAAGATGTCGACCAAG 721  
 |||||  
 Db 2754 CAGAATGAAAAAGAAATGCTGCAATGACACTTTGGAGAGAGATTAACCAAG 2813  
 QY 722 AATACATTTGCTGTTTGTGACTTCTTGAAATGCTGCTCATACATGCTGCTG 781  
 |||||  
 Db 2814 AATAGTGGGGCATATTTATCTTGCCCTC--AAATCATCACTTACTACTGTT 2870  
 QY 782 CCAGCTTACTGTTGATGCTGTTTCACTTCT 813  
 |||||  
 Db 2871 CAAATATTTGCTGTATGATGGGGTTATACATGT 2902  
 |||||  
 RESULT 14  
 AC105434/c 126105 bp DNA linear HTG 20-MAR-2004  
 LOCUS Magnaporthe grisea chromosome 7 clone 18114, \*\*\* SEQUENCING IN  
 DEFINITION PROGRESS \*\*\*, 2 ordered pieces.  
 AC105434  
 AC105434.2 GI:45597494  
 VERSION HTG, HTGS PHASE2.  
 KEYWORDS Magnaporthe grisea (anamorph: Pyricularia grisea)  
 SOURCE Magnaporthe grisea  
 ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.  
 1 (bases 1 to 126105)  
 REFERENCE 1  
 AUTHORS Thom, M.R., Mitchell, T., Brown, D.E., Diener, S., Taro, A., Pan, H. and  
 Dean, R.A.  
 TITLE The sequence of *Magnaporthe grisea* chromosome 7  
 JOURNAL Unpublished  
 2 (bases 1 to 126105)  
 REFERENCE 2  
 AUTHORS Dean, R.A., Dr, Mitchell, T., Dr, Thom, M. Dr and Brown, D.E.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-JAN-2002) Plant Pathology - Fungal Genomics  
 Laboratory, North Carolina State University, 840 Main Campus Drive,  
 Raleigh, NC 27606, USA  
 3 (bases 1 to 126105)  
 REFERENCE 3  
 AUTHORS Thom, M.R., Mitchell, T., Brown, D.E., Diener, S., Taro, A., Pan, H. and  
 Dean, R.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-MAR-2004) Center for Integrated Fungal Research,

## COMMENT

North Carolina State University, Plant Pathology - 840 Main Campus  
Dr, Raleigh, NC 27606, USA

On Mar 20, 2004 this sequence version replaced gi:18071329.

- \* NOTE: This is a 'working draft' sequence. It currently
- \* consists of 2 contigs. Gaps between the contigs
- \* are represented as runs of N. The order of the pieces
- \* is believed to be correct as given, however the sizes
- \* of the gaps between them are based on estimates that have
- \* provided by the submitor.
- \* This sequence will be replaced
- \* by the finished sequence as soon as it is available and
- \* the accession number will be preserved.
- \* 1 86195: contig of 86195 bp in length
- \* 86196 86395: gap of unknown length
- \* 86396 126105: contig of 39710 bp in length.

## FEATURES

source

1..126105 "MagnaPorte grisea"  
/organism="MagnaPorte grisea"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:148305"  
/chromosome="7"  
/clone="18114"

## ORIGIN

Query Match 10.6%; Score 86.8; DB 2; Length 126105;  
Best Local Similarity 53.6%; Pred. No. 3.3e-11;  
Matches 181; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 464 TTGTTTGAATGTTTATGTCGTGCGATTTGTCAAGATCTTCAAAACCAAGTTGT 523  
DB 56680 TGGTCTTCATCGCAGACATGAGGGCGCATGTCACGTCGCGACGCGAGGCCCT 56621  
QY 524 ACAACATGTCACAGCTGGTGTATCATTTGGCTTAAGACTTTGGCTTGATGGGCTA 583  
DB 56620 ACAACAGCTCAAGGCCCGCTGCCCACTGGCCCGACCTGGCCGCTGAGTGGCCG 56561  
QY 584 AGTACACATCAGAGTTAATCTTTAAACCAAGTTACATCTACGGTCTTGACCAAGA 643  
DB 56560 AGCGCCGATCGCGCTCACTGATCTCCCGGCTACATGCTCACCGCCCTGACCAAGA 56501  
QY 644 AGTTATCATGTTAAGCAAGATTTGTACACAGATGATCTGTGATATCCCAACAA 703  
DB 56500 AGATTCACGACGACCAAGCCGACCTCAAGAAAGAGACGACGACCTCATGCCCAAGGGA 56441  
QY 704 GATGTCGACCAAGCAATACATTTGCTGTTTGTACTTCTTCTGATCTGCTG 763  
DB 56440 GGATGGCCCCCAGAGGACGATGAGTCCGCTGACCTTCTGCTGTGGAGCGTCCG 56381  
QY 764 CTTTATACATCTAGTGTGCCAGCTTACTGCTTGATGCT 801  
DB 56380 TCTACGTCACCGCGCTGACATCCGTGATGAGCGGT 56343

## RESULT 15

ARS47359

LOCUS ARS47359 927 bp DNA linear PAT 08-OCT-2004

DEFINITION Sequence 2490 from patent US 6747137.

ACCESSION ARS47359

VERSION ARS47359.1 GI:53940534

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

## ORIGIN

Query Match 10.6%; Score 86.6; DB 6; Length 927;  
Best Local Similarity 52.0%; Pred. No. 4.3e-11;  
Matches 328; Conservative 0; Mismatches 279; Indels 24; Gaps 5;

QY 182 CTGAAGAAATGGAAGTTGAAGAAGTTCCAAAGATGGTTCAATGCTGTGATTTCTG 241  
DB 308 CTGAATATTTTAAACCAAGCTTATGGGCTTAAATCCAAAGCATACAAATGATATGTACTG 367  
QY 242 ATTGATACGCTTACAAAGGTTTGTCTCAAGTTGTAAAGATTGTGTAGTGGCAT 301  
DB 368 ATTTCAAAGATTTAAAGTTTGTCAAAATTTGAATTTGATTTGCTGATTTGTA 427  
QY 302 TGCATTTGTTTAAACAGCTGGTACTG---TGAACCTTCCATGTGAAGATTACCGAG 358  
DB 428 TCTTTGTTGCCAATCTGTGTGTGTGCTGACCGAAGGGCCGGAATGATGCAAGGAG 487  
QY 359 CCAAGACGCTGAGAAGATGTTAACTTTGGTGGTCTTTGTATGTTTCTCAG 418  
DB 488 TCGACAAATGGAACAAAGTTGTGATGTTGAATTTAAACAGTGTATATTTGTCTCATG 547  
QY 419 CTTTGTCAAGCATGTATCAAGAAAGATTCAGAGGCTTCTGTGTTTGTATGTT 478  
DB 548 TTGTTGTCATTTTCAAGAAAAGGTA-----AAGGTCATTCATTTTCACTGCCA 601  
QY 479 CTATGTCGTGTCATTTGTCAACGATCTCAAAACCAAGTTGTCTACACATGTTCAAG 538  
DB 602 GTATGTCGGCTTCAATTTGTATATGTCACAAATTCAGACAGCTTACAAAGCTGTAAAG 661  
QY 539 CTGTTGTTATTCATTTGGCTTAAGACTTTGGCTTTGGAATGGGCTTAAGTACATCAG 598  
DB 662 CTGGGGTCAACATTTGTCAAATCATTTGAGTGAATGGGACCAATTTGCT--AGAG 718  
QY 599 TTAATTTCTTAAACCAAGTTTACATCTACGCTTGTGACCAAGAAATGTATCAATGTA 658  
DB 719 TCAATTTGTTTCTTCAAGTTATCATGCTACTCATTTTGAATTTCTG----- 769  
QY 659 ACGAAGAAATGTACAAAGATGATCTGTGATCTCCCAACAAAGAAATGTCCGAAACCA 718  
DB 770 ATCCGAGATGTCAAGATTAATGTTGCAACTTACACGACTTGATGAGAAGCAAAACCA 829  
QY 719 AGGAATACATTTGCTGCTTTTGTATCTGCTTTTGTGAATCTGCTGCTTACACTACTG 778  
DB 830 GAGAGCTTTGTTGCTTACTTATTTTGGCTTCCGA---TGCTGACATCTTATTAACACTG 886  
QY 779 GTGCCAGCTTACTGTTGATGTTGTTTAC 809  
DB 887 GTGCTGATCTTGTGTGATGGTTACAC 917

Search completed: March 2, 2005, 15:57:30  
Job time : 3861 secs